Some Statistical Learning Methods for Procuring Individualized Treatment Rules in Personalized Healthcare

by

Yiwang Zhou

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biostatistics)
in The University of Michigan
2021

Doctoral Committee:

Professor Peter X.K. Song, Chair
Professor Karen E. Peterson
Assistant Professor Zhenke Wu
Professor Min Zhang

Yiwang Zhou

yiwangz@umich.edu

ORCID iD: 0000-0002-8023-205X

ACKNOWLEDGEMENTS

Throughout the writing of this dissertation, I have received a great deal of support and assistance from my supervisor, committee members, collaborators, as well as my families and friends.

I would first like to thank my supervisor, Professor Peter Song, who provides me with a lot of guidance and advice for my dissertation research. His expertise and insights in biostatistical research help me greatly in the formulation of my research questions and the development of my methodologies in the field of personalized healthcare. I will always remember the time talking with him about my research progress and the manuscripts revised by him with full of comments. His passion for scientific research inspires me to start a career in academia. I hope that I can become a successful researcher like him in the future.

I would also like to acknowledge my dissertation committee, including Professor Karen Peterson, Professor Min Zhang, and Assistant Professor Zhenke Wu, for their advice and suggestions on my dissertation research. Specifically, I want to thank Karen for providing me with the calcium supplementation trial data as a real data application in my second and third chapters. Her inputs of the background knowledge in nutritional sciences help me understand the data better for the derivation of individualized treatment rules in precision nutrition.

My dissertation research has been motivated and benefited significantly from several collaborative projects. I want to thank Dr. Haoda Fu from Eli Lilly for providing me with the diabetes clinical trial data for illustrating the performance of my pro-

posed method in Chapter II. I am grateful to Professor Ivo Dinov for giving me guidance and advice on a project studying the neuroimaging biomarkers in the UK Biobank. Finally, I am thankful to the whole ELEMENT DMMC group for sharing ideas and discussions on data analyses for the ELEMENT cohort study. All these experiences serve as the foundation for me to make progress step by step and will continue to nourish me until I grow up as an independent researcher in biostatistics.

Last but not least, I want to express my sincerest gratitude to my families and friends. Thanks to my parents, Yongyuan Zhou and Huiyu Lu, for supporting me throughout this entire time of Ph.D. study. Their support is my greatest encouragement on the journey of knowledge pursuing. Thanks to my husband, Jiyuan Yang, for always standing by my side when I need him and for giving me plenty of space when I want to be alone. His endless love, support, and understanding make me realize that I always have a harbor to rely on. Thanks to Ming Tang, Yingchao Zhong, Sai Chen, Huayun Hou, Lu Xia, and all my other friends for supporting me through the darkest time in my life so far. Thanks all friends for taking care of me and accompanying me. Without their help, I will not be able to get out of the most difficult time and finish my doctoral dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Precision health has gained increasing attention in recent years to deliver individualized healthcare to patients. One central task of precision health is to establish individualized treatment rules (ITRs) for patients through their heterogeneous responses in that tailored interventions are undertaken to maximize therapeutic effects. Although various methods have been proposed to estimate optimal ITRs, specific methodological challenges remain to be addressed, such as validating an existing ITR, procuring an ITR with multiple views and missing clinical data. In this dissertation, we develop new statistical learning methods to extend and improve the current methodologies used to derive ITRs.

Chapter II focuses on a task of clinical importance: utilizing novel candidate biomarkers to update an existing ITR to improve clinical benefits. I propose a new statistical framework, termed net benefit index (NBI), to quantify added values of candidate biomarkers when they are included in an existing ITR. NBI works by assessing a contrast between the resulting gain-and-loss of benefits when a biomarker enters an ITR to reallocate patients in treatments. To account for sampling uncertainty in gauging the contribution of a biomarker, I propose an NBI-based test statistic whose empirical null distribution is constructed by treatment-stratified permutation. I apply this NBI method to a randomized clinical trial on patients with type 2 diabetes to select useful biomarkers to improve an existing ITR for assigning pioglitazone or gliclazide to diabetic patients. As a result, the updated ITR helps maximally reduce

fasting plasma glucose concentration over the 52 weeks of treatment.

Chapter III concerns the development of a multi-view statistical learning approach to procuring ITR for precision nutrition, in the hope to maximize benefits of subjects taking nutritional supplementations. A key issue pertains to the availability of multiple endpoints for health benefits with different degrees of clinical relevance. I propose a new statistical learning method, termed synergistic self-learning (SS-learning), to address two major methodological challenges in ITR derivation, including heterogeneous multivariate outcomes and complex missing data patterns. I show both theoretically and numerically that SS-learning can effectively synergize heterogeneous benefit variables from multiple training data sources. I apply SS-learning to a calcium supplementation trial and identify several important predictors critical in ITR estimation. The resulting ITR would lead to a higher expected reduction in prenatal lead exposure, measured by infant blood lead concentration at birth, should it be implemented to pregnant women in the population.

Motivated from the calcium supplementation trial, Chapter IV centers at a further extension of SS-learning to longitudinal self-learning (LS-learning), through which I aim to establish an effective ITR using longitudinal benefit outcomes. The endpoint of the trial is child's blood lead concentration at age of 3 years old, so the resulting ITR can guide pregnant women for their daily calcium intake to maximize the long-term reduction of lead exposure by their children. I propose a temporally weighted self-learning paradigm in LS-learning, which enables to synergize auto-correlated training data sources with repeated measurements. I also develop a new tuning procedure by introducing a scaled tuning method, in which tuning parameters are governed by a certain function of observational time. Applying LS-learning to the calcium supplement trial, we establish an ITR from longitudinal benefit outcomes that contains several important predictors to reduce child's long-term lead exposure at age 3 should this ITR be implemented to the targeted pediatric population.

CHAPTER I

Introduction

Precision health, including topics of individualized medicine and individualized diet, has received increasing attention in recent years (*Council et al.*, 2011; *Collins and Varmus*, 2015). Instead of a one-drug-fits-all paradigm, precision health is an innovative framework, in which healthcare is customized with personal-level medical decisions, treatments, practices, and products. It proposes tailored medical practice to a subgroup of patients by individualized rules that take into account individual variability, for example, genes, environment, and lifestyle (*Yau*, 2019; *Niculescu et al.*, 2019; *Lu et al.*, 2014; *Jones et al.*, 2019). One major objective of personalized healthcare is to maximize therapeutic benefits for patients, who may experience heterogeneous responses to different types of treatments, by assigning them to a "right" treatment instead of the one based on the average treatment benefit for the entire target population.

One central analytic task in precision health is to establish individualized treatment rules (ITRs) that enable us to guide the assignment of treatments to individual patients according to their personal characteristics. Various approaches have been proposed in the literature to construct ITRs, including Q-learning (*Watkins*, 1989), A-learning (*Murphy*, 2003), outcome weighted learning (OWL) (*Zhao et al.*, 2012), residual weighted learning (RWL) (*Zhou et al.*, 2017), just to name a few. Q-learning

and A-learning have been shown to estimate the optimal ITR effectively under correctly specified regression models. However, both these approaches fail to give stable results when the regression model used for ITR derivation is misspecified. This shortcoming is rooted in the fact that a regression-based method focuses on the prediction of benefit outcomes, which is not the direct solution to the derivation of ITR in precision medicine, which aims at maximizing the expected benefits of interest (*Murphy*, 2005). In order to derive ITRs under the real goal of precision medicine, different methodologies that directly optimize the population-level expected outcomes have been proposed. *Zhao et al.* (2012) proposes OWL and shows that estimating the optimal ITR is equivalent to a classification problem of treatment groups, in which subjects in different groups are weighted proportionally to their observed clinical benefits. In OWL, support vector machine (SVM) is invoked to estimate the optimal ITR. Later, (*Zhou et al.*, 2017) proposes RWL, which shows to be able to improve the finite sample performance of OWL by weighting individual subjects with their residuals obtained by a regression fit of benefit outcomes on clinical covariates with no treatment variable.

Although some approaches developed for ITR derivation become widely used, specific research questions about applying precision medicine in clinical practice remain to be addressed, such as validating an existing ITR, procuring ITR with multiple views, and dealing with missing clinical data. Motivated by these practical challenges in clinical studies, this dissertation plans to develop a few new statistical learning methods to extend the techniques currently used for ITR derivation. In particular, we focus on the method of OWL in the establishment of ITRs due to its stable performance and easy applications for clinical settings.

Our first project in Chapter II considers the problem of utilizing promising candidate biomarkers to update the existing ITRs in order to achieve better clinical benefits for patients. It is motivated by a randomized clinical trial comparing the therapeutic

effects of two medications, pioglitazone and gliclazide, in treating patients with type 2 diabetes (*Charbonnel et al.*, 2005). The outcome of interest is the reduction rate of fasting plasma glucose (FPG) over the 52 weeks of treatment. An existing ITR was established involving age, body bass index (BMI) and baseline FPG, with the first two being common demographic variables and the last one being the baseline reference of the individual FPG level. With some new candidate biomarkers available, like Hemoglobin A1c (HbA1c), fasting insulin, and total cholesterol, we are interested in figuring out which biomarkers may be included into and update the existing ITR to achieve better clinical benefits if the new decision rule was implemented for the whole population. In order to select useful biomarkers, in Section 2.3 we propose, examine and illustrate a new statistical framework, termed *net benefit index (NBI)*, to analytically and numerically quantify added values of candidate biomarkers when they are included in the existing ITR. NBI works by quantifying a contrast between the resulting gain-and-loss of treatment benefits when a biomarker enters an existing ITR to reallocate patients in treatments. To account for sampling uncertainty in assessing the contribution of a biomarker, we propose an NBI-based test statistic for evaluating the significant improvement over the existing ITR, where the empirical null distribution is constructed via the method of stratified permutation by treatment arms. Simulation experiments in Section 2.4 show that the proposed NBI method achieves high sensitivity and specificity in selecting the useful biomarkers to improve ITR, resulting in a cost-effective selection of useful biomarkers and a better ITR with higher prediction accuracy in terms of the underlying correct treatment for patients. Applying NBI to the motivating diabetes trial, in Section 2.5 we find that baseline fasting insulin is an important biomarker that leads to an improvement over the existing ITR based only on patient's age, BMI and baseline FPG to reduce the FPG concentration over a period of 52-weeks' treatment.

Chapter III and Chapter IV focus on the estimation of ITRs for precision nutri-

tion. Different from precision medicine, precision nutrition aims to derive the optimal ITR to maximize the benefits for subjects in taking nutritional supplementations, such as vitamins, calcium, and hormone therapy. Both chapters are motivated by a randomized clinical trial in the third cohort study of the Early Life Exposure in Mexico to ENvironmental Toxicants (ELEMENT) Project (*Perng et al.*, 2019). The central objective of this clinical trial is to study the effect of daily intake of calcium supplementation by pregnant women in reducing infant blood lead concentration. It has been demonstrated in the literature that at the population level, dietary calcium supplementation can reduce maternal lead concentration in blood and thus alleviate lead exposure to infants through blood transfusion and breastfeeding (*Ettinger et al.*, 2007; *Janakiraman et al.*, 2003; *Gulson et al.*, 2004; *Téllez-Rojo et al.*, 2006a). Our interest lies in establishing an ITR that would guide pregnant women for their daily intake of calcium supplementation, which presents a useful nutritional intervention offered in nutritional counselling. This precision nutritional rule delivers guidelines under tailored individual dietary patterns, lifestyle and living conditions. The core problem in these two chapters pertains to the existence of multiple benefit outcomes, arising from practical issues related to different degrees of clinical relevance, measurement qualities, and missing data. They all present great technical challenges to the derivation of ITRs.

Chapter III presents a new method, termed synergistic self-learning (SS-learning), to address two major challenges for ITR derivation in the presence of multiple clinical outcomes, including heterogeneous multidimensional outcomes and complex missing data patterns. SS-learning can effectively synergize heterogeneous features of multiple training data sources in the derivation of ITRs. SS-learning works in a weighted self-learning paradigm, in which individual outcomes of different clinical relevance are incorporated separately as weights in SVM, and the estimated subgroup labels from low quality data sources are iteratively calibrated with information from high

quality data sources. With the introduction of additional tuning parameters, the relative contributions from respectively low- and high-quality data sources are tuned in a greedy way according to a criterion through minimization of the sum of squared errors (SSE) of the predicted benefits. We prove the algorithmic convergence of the proposed SS-learning method. We apply SS-learning to analyze the ELEMENT calcium supplementation trial and identify several important predictors in the construction of an optimal ITR. This resulting ITR would give a larger reduction of maternal lead exposure to infants at birth during pregnancy should it be implemented to the population of pregnant women. The performance of SS-learning is also examined by comprehensive simulation studies, showing that SS-learning outperforms the standard OWL method in ITR derivation with respect to the prediction accuracy of the underlying correct treatment assignment as well as the estimated benefit of lead exposure reduction at population level.

Chapter IV further extends SS-learning to longitudinal self-learning (LS-learning) to accommodate longitudinal benefit outcomes with missing data. In real-world clinical trial practice, missing data are pervasive, which may occur in complicated ways. The mechanism of missing completely at random (MCAR) or that of missing at random (MAR) may not be the case in reality. Therefore, a method that can deal with missing data under the mechanism of not missing at random (NMAR) is needed. For ITR derivation in this chapter, we follow the idea of pattern mixture models (*Little*, 2008) to first stratify subjects based on their subject-specific endpoint measurements, and then integrate the longitudinal outcomes measured at different time points (terms as surrogate endpoint). LS-learning assumes the primary endpoint benefit outcome being the one measured at the last longitudinal visit, while surrogate benefit outcomes are those measured longitudinally prior to the last endpoint. In the presence of missing endpoint benefit outcomes, we utilize surrogate endpoints in ITR derivation. A technical difficulty of ITR derivation is that there may exist many surrogate

benefit outcomes, leading to a case of many tuning parameters in the application of LS-learning. This results in a considerable cost of computation time on parameter tuning if the greedy tuning method proposed in Chapter III were employed. To overcome this computational challenge, we introduce a scaled tuning procedure that works by specifying the tuning parameters as a systematic function of the observational time. In this way, we can significantly reduce the total number of parameters to be tuned to speed up computation. Simulation experiments illustrate that LS-learning with scaled tuning outperforms the standard OWL in ITR derivation with respect to the prediction accuracy for the underlying correct treatment assignment. Applying LS-learning with the scaled tuning scheme to the calcium supplementation trial, we found several important predictors, including dietary intake of CA, fiber, Fe and Zn, in the derivation of ITR. The estimated ITR results in the highest long-term reduction of maternal lead exposure for infants at age of 3 years old.

CHAPTER II

Net Benefit Index: Assessing the Influence of a Biomarker for Individualized Treatment Rules

2.1 Introduction

Utility of the newly discovered biomarkers, such as omics-markers, from basic sciences to facilitate better and more cost-effective clinical practice is of critical importance in translational medicine. In connection to the emerging field of personalized medicine, one central task is to update the existing individualized treatment rules (ITRs) using new biomarkers with the aim to receive better clinical benefit. A noticeable shortcoming in the current statistical literature of personalized medicine is a lack of such methods to evaluate the significance of individual biomarkers in improving the existing ITRs. Perhaps this has been regarded as a small mathematical problem, but such methodological need is not easy to be addressed appropriately given many practical constraints involved, such as clinical benefit and medical cost associated with the inclusion of such biomarkers in daily clinical practice. This paper is intended to fill in this technical gap with a specific objective of developing a new statistical procedure to assess the usefulness of a biomarker in the context of personalized medicine.

Motivated from a randomized clinical trial comparing two drugs, i.e. pioglitazone

and gliclazide, on treating Type 2 diabetic patients, we propose, examine and illustrate a new statistical framework, termed as *net benefit index (NBI)*, to analytically and numerically quantify added values of candidate biomarkers when they are used to update an existing ITR. We consider an existing ITR involving age, body mass index (BMI) and baseline fasting plasma glucose (FPG) to maximize the reduction of FPG after 52 weeks of treatment. With several new variables like Hemoglobin A1c (HbA1c), fasting insulin, etc., we want to evaluate their added values, and decide which one or ones can significantly improve the existing ITR. Added value of a promising biomarker is gauged by a contrast between the resulting gain and loss of clinical benefits from reallocations of treatments through a revised ITR with the utility of this new biomarker. Reallocation is optimized by outcome weighted learning (OWL) (*Zhao et al.*, 2012). In addition, we consider an NBI-based test for significance of a certain added value in which the empirical null distribution is generated via stratified permutation by treatment arms. Through this procedure, biomarkers that significantly improve an existing ITR are sequentially selected to revise current allocation rules, where the biomarker selection is controlled under false discovery rate (FDR) to avoid the issue of overfitting. Note that overfitting in terms of the number of biomarkers is practically unattractive due to higher costs and longer time spent on collecting samples that essentially produce redundant information, which may undermine the accuracy and interpretation of an estimated ITR.

While NBI allows to evaluate the contribution of new biomarkers under controlled FDR, it avoids some other problems known in existing biomarker selection techniques in the context of personalized medicine. In the field of decision-making, variable selection should target primarily at prescriptive variables that help prescribe the optimal action, instead of the predictive variables that reduce the variability and increase the accuracy of an estimator. A prescriptive variable has to have a qualitative interaction with the treatment (*Gunter et al.*, 2011). A variable is said to qualitatively inter-

act with the treatment if there exists at least two distinct, non-empty sets within the space of the variable, for which the treatment arms that maximize the expected clinical benefit are distinct (*Gunter et al.*, 2011). *Qian and Murphy* (2011) proposes a two-stage Q-learning (Q denoting "quality") (*Watkins*, 1989) procedure that employs the l_1-penalty for variable selection to estimate an optimal ITR. *Lu et al.* (2013) develops a penalized regression framework, known as a kind of A-learning (A denoting "advantage") (*Murphy*, 2003) that allows to simultaneously estimate an optimal ITR and to select an important variables. However, neither of these two methods specifically targets at the selection of prescriptive variables. *Gunter et al.* (2011) proposes two variable-ranking criteria, U-score and S-score, for variable selection via qualitative interactions. But one limitation of these criteria is the ignorance of the correlations between the variables. To overcome this issue, *Fan et al.* (2016) develops a sequential advantage selection (SAS) method based on a modified version of S-score. SAS sequentially evaluates additional values of new variables via qualitative interactions, so that it can avoid identifying any variables marginally important but jointly unimportant. However, SAS lacks its direct relevance in clinical practice as it does not directly optimize treatment benefit objective for ITR, instead building models with sequentially added interaction terms under a statistical criterion of mean squared error. Different from these existing methods, NBI has the following advantages: (i) NBI directly optimizes treatment grouping labels to maximize the expected clinical benefit; (ii) NBI selects important prescriptive variables beneficial for treatment allocation; (iii) NBI is naturally applicable for non-linear decision rules due to the invocation of support vector machine (SVM); (iv) NBI sequentially selects biomarkers into an existing ITR through FDR control.

The remainder of the chapter is organized as follows. Section 2.2 introduces the motivating diabetes clinical trial, followed by the framework of NBI for biomarker assessment in Section 2.3. Section 2.4 evaluates the proposed NBI test through sim-

ulation experiments. The NBI method is illustrated by the motivating clinical trial in Section 2.5. Section 2.6 contains some concluding remarks. Detailed introduction of OWL and some additional simulation results are available in Appendix A-D.

2.2 Application: A Diabetes Clinical Trial

This is a randomized control and double-blind trial that aims to compare the therapeutic effects of pioglitazone and gliclazide in treating patients with Type 2 diabetes. Pioglitazone and gliclazide are two common oral medications with different therapeutic mechanisms for the treatment of Type 2 diabetic patients. A total of 1,270 patients with Type 2 diabetes were recruited into the trial. All the eligible patients were randomized to a 52-week treatment period. The outcomes of interest is the change of FPG between the last post-treatment measurement and baseline. FPG was measured repeatedly at baseline and at week 4, 8, 12 up to 52. Other variables measured at baseline included age, BMI, HbA1c, fasting insulin, high-density lipoproteins (HDL), low-density lipoproteins (LDL), aspartate transferases (AST), alanine transferases (ALT), total cholesterol, triglycerides, creatinine, and gamma-glutamyl transferase (GGT). After deleting subjects with missing data, a sample of 830 patients remains available for analysis, with 424 assigned to pioglitazone and 406 assigned to gliclazide. Due to loss of follow-up, some of the patients had the last post-treatment measurement taken at week 32 or 42, resulting in a shorter period of treatment. *Charbonnel et al.* (2005) performs a non-inferiority test for the differential treatment effects between these two drugs on the reduction of FPG and illustrates a significantly greater mean reduction of FPG by pioglitazone (2.4 mmol/L) than by gliclazide (2.0 mmol/L), with a treatment difference of 0.4 mmol/L in favour of pioglitazone (95% CI 0.1 to 0.7 mmol/L). The comparison of FPG reduction rate illustrates that pioglitazone leads to a more effective FPG reduction (0.049 mmol/L/week) than gliclazide (0.038 mmol/L/week), with a treatment difference of 0.011 mmol/L/week

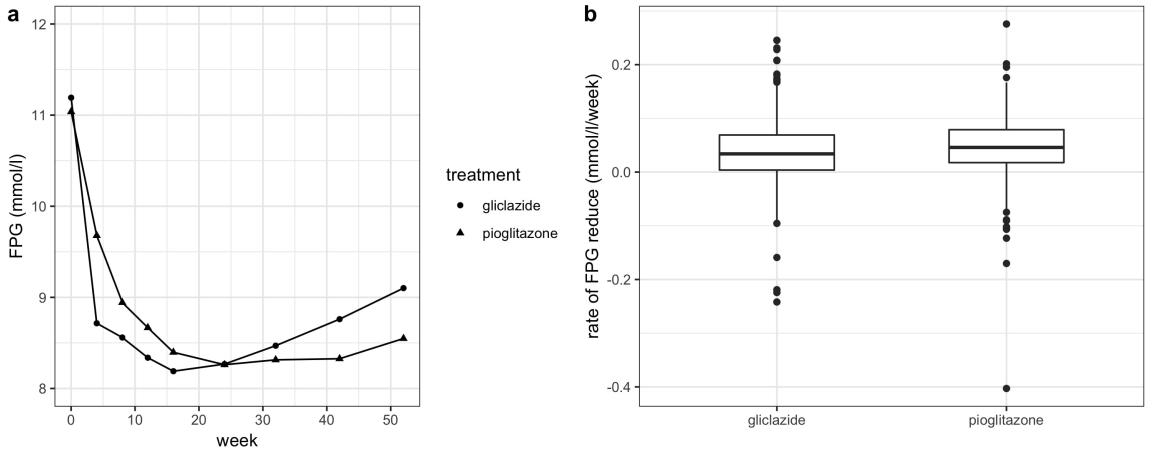

Figure 2.1: (a) Effects of pioglitazone and gliclazide on reducing fasting plasma glucose (FPG) during the 52-week period. (b) Average reduction rate of FPG by pioglitazone and gliclazide in the 52-week period.

(95% CI 0.003 to 0.019 mmol/L/week) (Figure 2.1).

The previous analysis finds a significant treatment difference on population average between pioglitazone and gliclazide in reducing FPG. However, for individual patients, some taking pioglitazone may receive little benefit, while some taking gliclazide may receive significant benefit. Given such heterogeneous responses to these drugs, an optimal ITR is deemed necessary to increase the benefit by shuffling patients in a systematic way to assign each patient to the "right" drug. This purpose of reallocation may be formulated and achieved with the aim to maximize the expected FPG reduction via a revised drug allocation rule. Consider a simple preliminary ITR that involves only baseline FPG, age and BMI, denoted by ITR(b.FPG, age, BMI). Age and BMI are two commonly used demographics for treatment assignment, while baseline FPG is a key clinical factor representing a personal reference level for the target endpoint. Among the available additional candidate biomarkers, we want to determine which ones may provide significant added values to improve ITR(b.FPG, age, BMI); if there are some, the expanded ITR is expected to provide higher treatment benefit than that given by the preliminary ITR.

2.3 Formulation

2.3.1 Outcome Weighted Learning and Optimal ITR

Consider a two-armed randomized clinical trial where each patient is randomly assigned a treatment $A \in \mathcal{A} = \{-1, 1\}$. $A = -1$ is the traditional treatment, say *gliclazide*. $A = 1$ is the new treatment, say *pioglitazone*. Complete randomization implies that the treatment allocation scheme is independent of patients' prognostic variables, denoted as $\boldsymbol{X} = (X_1, ..., X_d)^\intercal \in \mathcal{X} \subseteq \mathbb{R}^d$. Potential clinical benefit $B^*(A)$ is the outcome that would result if a patient were assigned to A. Since each patient takes only one treatment, the observed clinical benefit is given by $B = I(A = 1)B^*(1) + I(A = -1)B^*(-1)$, which is determined by A. The other potential clinical benefit is latent with no data captured. Suppose that B is bounded with a larger value of B being clinically more desirable. The primary aim of personalized medicine is to establish a decision rule $D(\boldsymbol{X})$, a mapping $\mathcal{X} \to \mathcal{A}$, that maximizes the expectation of the clinical benefit. The following three assumptions are typically required for computing the expectation of the clinical benefit (*Berkane*, 2012): a) Consistency assumption: $B = I(A = 1)B^*(1)+I(A = -1)B^*(-1)$; b) No unmeasured confounders assumption: $A \perp \{B^*(a)\}_{a\in\mathcal{A}} \mid \boldsymbol{X}$; c) Positivity assumption: $P\{P(A = a \mid \boldsymbol{X}) > 0\} = 1, \forall a \in \mathcal{A}$.

In this paper, the optimal ITR refers to the decision rule $D^*(\boldsymbol{X})$ that maximizes the expected clinical benefit $V(D) = E\left\{\frac{I(A=D(\boldsymbol{X}))}{P(A|\boldsymbol{X})}B\right\}$, which according to *Zhao et al.* (2012) may be formulated as a weighted classification problem that can be solved by SVM in the context of OWL (see details in Appendix B). The optimization problem of OWL has very relevant interpretations to our definition of NBI. For patients with large observed benefits, the optimality encourages to allocate the same treatment type as the one previously assigned. Conversely, for those receiving small observed benefits, the optimality tends to assign the alternative treatment type. Note that in

the diabetes trial, the actually implemented allocation is complete randomization, i.e. $P(A \mid \boldsymbol{X}) = P(A) = 0.5$. But this might not give the best personalized drug allocation rule as randomization primarily aims to control confounding not to maximize treatment benefit. Clearly, the optimal decision rule $D^*(\boldsymbol{X})$ will give a higher overall clinical benefit in comparing to completely randomized trial.

2.3.2 Treatment Reallocation

Let $D(\boldsymbol{X})$ be a decision rule based on features $\boldsymbol{X}$. The optimization for $D(\boldsymbol{X})$ imposed by OWL encourages concordant treatment assignment on patient who receives clearly treatment benefit. In other words, with the invocation of OWL, patients who are assigned $D(\boldsymbol{X}) = A$ tend to have larger benefit than those who receive $D(\boldsymbol{X}) \neq A$. Denote an existing decision function by $f_e(\boldsymbol{X}_e)$ and the corresponding decision rule by $D_e(\boldsymbol{X}_e)$ based on the existing variables $\boldsymbol{X}_e$. Likewise, denote an updated decision function by $f_u(\boldsymbol{X}_e, \boldsymbol{X}_u)$ and the corresponding updated decision rule by $D_u(\boldsymbol{X}_e, \boldsymbol{X}_u)$ by involving new variables $\boldsymbol{X}_u$. Under decision rules $D_e(\boldsymbol{X}_e)$ and $D_u(\boldsymbol{X}_e, \boldsymbol{X}_u)$, there exist two subgroups of patients who receive the same treatment allocation, i.e. $D_e(\boldsymbol{X}_e) = A$ and $D_u(\boldsymbol{X}_e, \boldsymbol{X}_u) = A$, respectively. For those patients who are assigned the same treatment by $D_e(\boldsymbol{X}_e)$ and $D_u(\boldsymbol{X}_e, \boldsymbol{X}_u)$, the inclusion of a new biomarker does not lead to any benefit gain. Thus they should be excluded from the assessment of the difference caused by the biomarker. In other words, only those patients who are assigned a different treatment by $D_u(\boldsymbol{X}_e, \boldsymbol{X}_u)$ from that by $D_e(\boldsymbol{X}_e)$ should be used to define an effective amount of clinical benefit.

Figure 2.2 illustrates a simple example showing the different groups of subjects included in the calculation of $V(D)$ when new biomarker X_2 is used in the learning of an updated decision function $f(X_1, X_2)$ compared to $f(X_1)$. Subjects randomly assigned to $A = 1$ and $A = -1$ in the trial are denoted by circles and triangles. The sizes of circles and triangles reflect the magnitude of clinical benefit, with a bigger

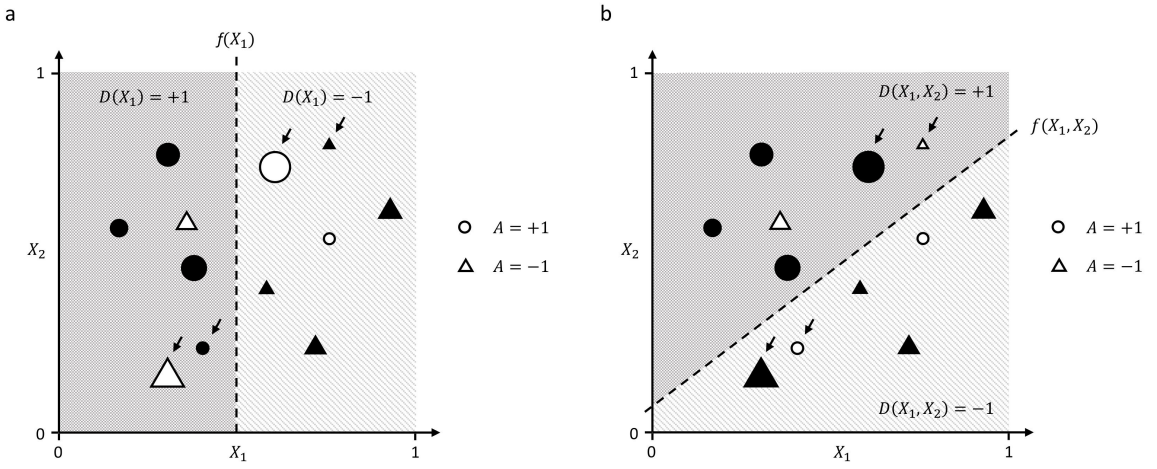

Figure 2.2: Changes of the subjects included in the calculation of $V(D)$ when new biomarker X_2 is included in the estimation of the decision function. Sizes of the circles and triangles reflect the magnitude of the clinical benefit B. Black circles and triangles are the subjects included in the calculation of $V(D)$ since they have $D(\boldsymbol{X}) = A$. Subjects in group "gain" and "loss" are pointed by arrows. (a) Decision function $f(X_1)$ estimated only on the existing biomarker X_1. (b) Decision function $f(X_1, X_2)$ estimated on X_1 and X_2.

size corresponding to a larger benefit. Only those subjects in black are effectively included in the calculation of $V(D)$ since they are assigned the concordant treatment $D(\boldsymbol{X}) = A$. By comparing the black circles and triangles in Figure 2.2(a) and 2.2(b), we can find that a circle and a triangle (pointed by arrows) are newly included in the calculation of $V(D)$, indicating that there is a gain as the consequence of reallocation by $f(X_1, X_2)$. At the same time, another circle and another triangle (pointed by arrows) are excluded from the calculation of $V(D)$, indicating that there is a loss. Since the newly included subjects have larger benefit (larger size), the gain exceeds the loss in the expected benefit value. Note that only subjects with $f(X_1) \times f(X_1, X_2) < 0$ will be included into the respective groups "gain" and "loss" since they have switching allocations from $D(\boldsymbol{X}_e)$ to $D(\boldsymbol{X}_e, \boldsymbol{X}_u)$. Technically, they are the ones responsible for a difference in the calculation of $V(D)$. It is conceptually appealing to quantify the contrast between "gain" and "loss" to understand the influence of a new biomarker for added value in personalized treatment allocation. We assume the following ethics

conditional on reallocation treatment.

Assumption II.1. *Let $D(\boldsymbol{X})$ be an allocation rule based on variable $\boldsymbol{X}$, and let $B(A \mid \boldsymbol{X})$ be the observed benefit when $D(\boldsymbol{X}) = A$. Suppose $\forall \epsilon > 0$, $\exists \delta(\epsilon) < \epsilon$, such that $P\{B(A^c \mid \boldsymbol{X}) \geq B(A \mid \boldsymbol{X}) \mid B(A \mid \boldsymbol{X}) < \epsilon\} \geq 1 - \delta(\epsilon)$, where A^c is the alternative treatment to A.*

Assumption II.1 implies that when the clinical benefit of a patient receiving treatment A tends to zero, with probability approaching to 1 there is no loss of benefit for allocating the alternative treatment A^c to the patient.

2.3.3 Net Benefit Index (NBI)

Let B_i be the observed benefit value for each patient, $i = 1, \ldots, n$. Denote $\mathcal{S}_{\text{gain}}$ as the sample of patients in group "gain", and $\mathcal{S}_{\text{loss}}$ as the sample of patients in group "loss". A net benefit index for a new biomarker X_u is defined as follows:

Definition II.1. ***(Net Benefit Index).***

$$NBI(X_u) = \frac{\sum_{i \in \mathcal{S}_{gain}} B_i / P(A_i \mid \boldsymbol{X}_i)}{\sum_{i \in \mathcal{S}_{gain}} 1 / P(A_i \mid \boldsymbol{X}_i)} - \frac{\sum_{i \in \mathcal{S}_{loss}} B_i / P(A_i \mid \boldsymbol{X}_i)}{\sum_{i \in \mathcal{S}_{loss}} 1 / P(A_i \mid \boldsymbol{X}_i)}. \tag{2.1}$$

Remark II.1. *In the case of a randomized clinical trial, propensity $P(A_i \mid \boldsymbol{X}_i) = P(A_i) \equiv 0.5$ for $i = 1, \ldots, n$. Let $n_{gain} = |\mathcal{S}_{gain}|$ and $n_{loss} = |\mathcal{S}_{loss}|$, NBI becomes:*

$$NBI = \sum_{i \in \mathcal{S}_{gain}} B_i / n_{gain} - \sum_{i \in \mathcal{S}_{loss}} B_i / n_{loss} = \overline{B}_{\mathcal{S}_{gain}} - \overline{B}_{\mathcal{S}_{loss}}, \tag{2.2}$$

which is actually the difference of the average observed benefits of $\mathcal{S}_{\text{gain}}$ and $\mathcal{S}_{\text{loss}}$. Clearly, NBI > 0 suggests that a new biomarker is potentially valuable to improve the existing ITR.

Remark II.2. *To account for sampling variability, we propose a standardized NBI as: standardized-NBI*$(X_u) = (\overline{B}_{\mathcal{S}_{gain}} - \overline{B}_{\mathcal{S}_{loss}})/\sqrt{\frac{s^2_{gain}}{n_{gain}} + \frac{s^2_{loss}}{n_{loss}}}$*, where* s^2_{gain} *and* s^2_{loss} *are sample variances of observed benefits for* $\mathcal{S}_{gain}$ *and* $\mathcal{S}_{loss}$*, respectively.*

The calculation of NBI and standardized NBI for X_u is given by Algorithm II.1. Note that the minimal sample size, $n_{\text{gain}} \geq 5$ and $n_{\text{loss}} \geq 5$, is required in the calculation of standardized NBI to have numerical stability.

Algorithm II.1 Calculation of standardized-NBI(X_u)

1: Establish models $f_e(\mathbf{X}_e)$ and $f_u(\mathbf{X}_e, X_u)$ using OWL on a training dataset.
2: Get classifications $D_e(\mathbf{X}_e)$ and $D_u(\mathbf{X}_e, X_u)$ for subjects in a NBI evaluation dataset.
3: Characterize samples $\mathcal{S}_{\text{gain}}$ and $\mathcal{S}_{\text{loss}}$ by comparing $D_e(\mathbf{X}_e)$ and $D_u(\mathbf{X}_e, X_u)$.
4: **if** $n_{\text{gain}} \geq 5$ and $n_{\text{loss}} \geq 5$ **then**
5: Calculate standardized-NBI(X_u) based on the values of B in $\mathcal{S}_{\text{gain}}$ and $\mathcal{S}_{\text{loss}}$.
6: **else** Set standardized-NBI(X_u)=0.

2.3.4 Test for Significant NBI

NBI> 0 is only suggestive subjective to sampling uncertainty, which may further be made rigorous by hypothesis testing. For a practical point of view, we hypothesized that $D_u(\boldsymbol{X}_e, \boldsymbol{X}_u)$ should not be inferior to $D_e(\boldsymbol{X}_e)$. Thus, we consider the following hypotheses: H_0: the new biomarker does not improve ITR; against H_a: the new biomarker improves ITR. Let μ_{gain} and μ_{loss} be the expected benefits in the "gain" and the "loss" population respectively. The hypotheses can be stated as a two-sample mean comparison: $H_0 : \mu_{\text{gain}} = \mu_{\text{loss}}$; against $H_a : \mu_{\text{gain}} > \mu_{\text{loss}}$. We will apply the standardized NBI to perform the hypothesis of two sample mean comparison. However, it is difficult to derive the distribution of standardized NBI. Therefore, we invoke the empirical null distribution of the standardized NBI to generate the p-values.

Remark II.3. *Different from the standard two-sample test, here* $\mathcal{S}_{gain}$ *and* $\mathcal{S}_{loss}$ *are random sets generated by a resulting optimal reallocation of treatments under a com-*

mon overall optimal benefit objective function. Thus, there exists a certain shared action in group membership labeling, which leads to a dependence between these two sets. However, when conditional on the memberships of $\mathcal{S}_{gain}$ and $\mathcal{S}_{loss}$, we have the conditional independence, which leads to a standard unequal variance two-sample t-statistic, $t \mid \mathcal{S}_{gain}, \mathcal{S}_{loss} = \{\delta - (\overline{B}_{gain} - \overline{B}_{loss})\}/\sqrt{s^2_{gain}/n_{gain} + s^2_{loss}/n_{loss}} \sim t(\nu)$, where $\delta = \mu_{gain} - \mu_{loss}$ and $\nu = \nu(n_{gain}, n_{loss}, s^2_{gain}, s^2_{loss})$ is degrees of freedom. Clearly, the sizes of both sets, n_{gain} and n_{loss}, are random and correlated in ν. Since the labels in $\mathcal{S}_{gain}$ and $\mathcal{S}_{loss}$ have a rather complicated and unknown joint distribution, the marginal distribution of the t-statistic is not available to make inference. Thus, we invoke the empirical null distribution to obtain p-values.

To do so, we propose to create a null variable X_{null} via the means of permutation with projected residuals $r_i = X_{u,i} - \widehat{E}(X_{u,i} \mid \boldsymbol{X}_{e,i}), i = 1, \ldots, n$ as detailed in Algorithm II.2.

Algorithm II.2 Generation of empirical null distribution for standardized-NBI(X_u)

1: **if** $\boldsymbol{X}_e \neq Null$ **then**
2: Model $X_u = g(\boldsymbol{X}_e) + \epsilon$, where $g(\cdot)$ is a suitable function independent of A and B.
3: Get residuals $r_i = X_{u,i} - \hat{E}(X_{u,i} \mid \boldsymbol{X}_{e,i}), i = 1, \ldots, n$.
4: **else** Let $r_i = X_{u,i}$.
5: **for** $l = 1, \ldots, L$ **do** (L is the number of permutation replicates).
6: Permute the residuals conditional on A; get the permuted residuals $r^p_{l,i}, i = 1, \ldots, n$.
7: Values of $X_{\text{null},l}$ are generated as $X_{\text{null},l,i} = \widehat{E}(X_{u,i} \mid \mathbf{X}_{e,i}) + r^p_{l,i}, i = 1, \ldots, n$.
8: Calculate standardized-NBI($X_{\text{null},l}$) using Algorithm II.1.

Assumption II.2. *The new variable X_u can be expressed by an additive model $X_u = g(\boldsymbol{X}_e) + \epsilon$ of the existing variables $\boldsymbol{X}_e$ and the error term ϵ. Discussion of violations of Assumption II.2 is included in Section 2.6.*

Algorithm II.2 outputs the empirical null distribution of the standardized NBI, and the p-value is given as $p = \#\{\text{standardized-NBI}(X_{\text{null}}) > \text{standardized-NBI}(X_u)\}/L$.

The invocation of stratification by treatment arm in the permutation test is to retain the difference between the underlying distributions of the residuals across two treatment groups. Pooling the residual distributions together while performing permutation test would ruin the interaction effect between treatment and biomarkers. Since our major interest is to evaluate the added value of a biomarker when we have an existing ITR, we will focus our method on the situation when $\boldsymbol{X}_e \neq Null$ in the following simulation studies and real data analysis. Simulations results with $\boldsymbol{X}_e = Null$ are included in Table C.1 in Appendix C.

When there are several new variables under screening, say m, it is necessary to control FDR to ensure a balance of sensitivity and specificity. To proceed, we propose a forward selection method, Algorithm II.3, that sequentially adds the currently most significant variable with the smallest p-value into a current model at each step until no more variable to be added. The significant variables at each step are identified as those passing the FDR control through the Benjamini-Hochberg procedure.

Algorithm II.3 Sequential forward variable selection based on NBI test

1: Set $m = \dim(\boldsymbol{X}_u)$.
2: **while** $m > 0$ **do**
3: Get p_j for $X_{u,j}, j = 1, \ldots, m$ with the existing model involving $\boldsymbol{X}_e$ by NBI test.
4: Order $p_{(1)} \leq \cdots \leq p_{(m)}$, each corresponding to $H_{(j)}$: $\boldsymbol{X}_{u,(j)}$ does not improve ITR.
5: Find $j_{\max} = \max_j\{j : p_{(j)} \leq \frac{j}{m}q\}$, where $q \in (0, 1)$ is the chosen target FDR control.
6: **if** $j_{\max}$ exists **then** Update $\boldsymbol{X}_e = \{\boldsymbol{X}_e, X_{u,(1)}\}, \boldsymbol{X}_u = \boldsymbol{X}_u \setminus X_{u,(1)}$, and set $m = m - 1$.
7: **else** Stop.

2.4 Simulation Experiments

In this section, we conducted extensive simulations to evaluate the finite sample performance of the proposed NBI method.

2.4.1 Single-variable-based Decision Rule Evaluation

The first simulation concerns a setting in which an existing ITR consists of two variables X_1 and X_2, where $X_i \overset{i.i.d}{\sim} U(0,1), i = 1, 2$. Suppose that a new variable $X_u \sim U(0,1)$ becomes available, which is correlated with X_2, namely $\text{Corr}(X_2, X_u) = \rho$ with $\rho \in \{0.0, 0.2, 0.5, 0.8\}$. The following types of X_u are considered: (i) $X_u = X_3$, an important feature related to benefit B; (ii) $X_u = X_4$, a noise variable unrelated to B. Our goal is to assess the sensitivity (i.e. rate of detecting X_3) and specificity (i.e. rate of not detecting X_4) by the proposed NBI test. Allocation of treatment $A \in \{-1, 1\}$ is independent of $\boldsymbol{X}$ with $P(A \mid \boldsymbol{X}) = 0.5$. B is generated from a normal distribution with mean $\mu = 0.5 + X_1 + 2.0Af(\boldsymbol{X})$ and standard deviation $\sigma = 1.0$. Interaction term $Af(\boldsymbol{X})$ specifies a bimodal expected benefit that generates bifurcated benefit outcomes. We consider the following three scenarios of true decision function f:

1) (Linear) $f(\boldsymbol{X}) = 1 - X_1 + X_2 - 2X_3$;

2) (Binary) $f(\boldsymbol{X}) = 4\{I(X_1 > 0.1 \cap X_2 < 0.75 \cap X_3 > 0.25) - 0.5\}$;

3) (Nonlinear) $f(\boldsymbol{X}) = 2.5\{(X_1 - 0.5)^+ + (X_2 - 0.2)^+ - (X_3 - 0.1)^+\}$.

The sample size is set at $n = 800, 1000, 1200$. 5-fold cross-validation is used to determine the training dataset to learn f and the NBI evaluation dataset to assess X_u. We set type I error rate $\alpha = 0.05$. Simulation is repeated for 1000 times.

Table 2.1 summarizes the discovery rates of X_3 and X_4 across different f based on the proposed NBI method. It is shown that the sensitivity is high in detecting the useful variable X_3, and type I error has been well controlled for the noise variable X_4 at the nominal level 0.05. Table 2.2 reports the NBI values for X_3 and X_4. Aligned with the high sensitivity, the corresponding NBI(X_3) are all positive, implying that the inclusion of X_3 results in an improved ITR(X_1, X_2, X_3), in which more patients are assigned into their beneficial treatment arm in comparison to the previous ITR(X_1, X_2). In contrast, all the NBI(X_4) values are negative, indicating that the

Table 2.1: Discovery rates for X_3 and X_4 in the single-variable-based decision rule evaluation. (Discovery rate for X_4 equals 1-specificity.)

		$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
scenario	n	X_3	X_4	X_3	X_4	X_3	X_4	X_3	X_4
linear	800	0.987	0.047	0.991	0.056	0.984	0.057	0.931	0.053
	1000	0.998	0.054	0.997	0.043	0.992	0.051	0.958	0.055
	1200	0.999	0.059	1.000	0.051	0.998	0.052	0.974	0.053
binary	800	0.932	0.051	0.946	0.049	0.956	0.050	0.815	0.050
	1000	0.956	0.049	0.967	0.034	0.971	0.046	0.866	0.051
	1200	0.968	0.054	0.975	0.048	0.972	0.045	0.927	0.050
nonlinear	800	0.977	0.053	0.987	0.055	0.973	0.055	0.854	0.050
	1000	0.984	0.055	0.996	0.058	0.983	0.042	0.906	0.044
	1200	0.995	0.043	0.997	0.049	0.994	0.054	0.934	0.036

inclusion of X_4 results in a worse updated ITR(X_1, X_2, X_4) that assigns more patients into their non-beneficial treatment arm. When FDR is controlled, the chance of X_4 entering the updated ITR is slim, and the resulting decline in outcome of benefit is indeed ignorable.

2.4.2 Multiple-variable-based Decision Rule Evaluation

The second simulation uses the same setup of the base ITR(X_1, X_2) specified in Section 2.4.1. Now we consider multiple signal and noise candidate biomarkers $X_j \sim U(0,1), j = 3, \ldots, 12$, in which only X_3 and X_4 are signal biomarkers involved in the optimal ITR. The correlation structure of the variables is that $\text{Corr}(X_3, X_5) = \text{Corr}(X_4, X_6) = 0.5$, and $\text{Corr}(X_s, X_t) = 0.2, s, t \in \{7, \ldots, 12\}, s \neq t$. The mean parameter of benefit B is set as $\mu = 0.5 + X_1 + 2.0Af(X)$. We consider the following three scenarios of true decision function f:

4) (Linear) $f(\boldsymbol{X}) = 0.5(1 + X_1 + X_2 - 1.8X_3 - 2.2X_4)$;

5) (Binary) $f(\boldsymbol{X}) = 6\{I(X_1 > 0.12 \cap X_2 < 0.88 \cap X_3 > 0.2 \cap X_4 < 0.8) - 0.5\}$;

6) (Nonlinear) $f(\boldsymbol{X}) = (X_1 - 0.9)^+ - (X_2 - 0.78)^+ + (X_3 - 0.1)^+ - (X_4 - 0.22)^+$.

Table 2.2: NBI values for X_3 and X_4 in the single-variable-based decision rule evaluation.

		$\rho = 0.0$		$\rho = 0.2$	
scenario	n	X_3 mean (sd)	X_4 mean (sd)	X_3 mean (sd)	X_4 mean (sd)
linear	800	1.708 (0.547)	-0.150 (0.636)	1.724 (0.556)	-0.136 (0.653)
	1000	1.700 (0.517)	-0.139 (0.622)	1.693 (0.509)	-0.194 (0.611)
	1200	1.657 (0.445)	-0.161 (0.614)	1.655 (0.442)	-0.140 (0.611)
binary	800	3.139 (1.553)	-0.155 (1.922)	3.192 (1.489)	-0.217 (1.815)
	1000	3.090 (1.449)	-0.127 (1.827)	3.121 (1.382)	-0.160 (1.835)
	1200	3.075 (1.386)	-0.081 (1.779)	3.069 (1.287)	-0.046 (1.776)
nonlinear	800	1.793 (0.883)	-0.224 (0.888)	1.862 (0.804)	-0.265 (0.894)
	1000	1.793 (0.791)	-0.227 (0.883)	1.856 (0.740)	-0.293 (0.889)
	1200	1.770 (0.723)	-0.228 (0.827)	1.855 (0.695)	-0.277 (0.791)
		$\rho = 0.5$		$\rho = 0.8$	
scenario	n	X_3 mean (sd)	X_4 mean (sd)	X_3 mean (sd)	X_4 mean (sd)
linear	800	1.594 (0.628)	-0.127 (0.651)	1.125 (0.677)	-0.158 (0.688)
	1000	1.551 (0.552)	-0.148 (0.626)	1.154 (0.563)	-0.145 (0.626)
	1200	1.503 (0.487)	-0.171 (0.616)	1.152 (0.511)	-0.132 (0.618)
binary	800	3.329 (1.499)	-0.224 (1.855)	2.285 (1.888)	-0.208 (1.760)
	1000	3.172 (1.318)	-0.056 (1.751)	2.345 (1.685)	-0.154 (1.813)
	1200	3.252 (1.231)	-0.029 (1.756)	2.573 (1.622)	-0.139 (1.653)
nonlinear	800	1.791 (0.796)	-0.284 (0.903)	1.188 (0.832)	-0.264 (0.932)
	1000	1.801 (0.689)	-0.329 (0.873)	1.293 (0.743)	-0.343 (0.912)
	1200	1.823 (0.672)	-0.277 (0.841)	1.336 (0.692)	-0.310 (0.811)

We draw summary statistics under the FDR control set at 0.10. In addition to those basic performance measures considered in Section 2.4.1, we add a comparison of our NBI method on biomarker selection with SAS mentioned in Section 2.1 and riskRFE (*Dasgupta et al.*, 2019), a backward elimination method for variable selection developed for SVM.

Table 2.3 reports some summary statistics, including (i) size: the total number of selected biomarkers; (ii) true discovery rate (TDR): the number of correctly selected biomarkers over size; (iii) Matthews correlation coefficient (MCC): MCC $= \frac{(\mathrm{TP}\times\mathrm{TN}-\mathrm{FP}\times\mathrm{FN})}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}}$, where TP is true positive, TN is true negative, FP is false positive and FN is false negative; (iv) correct classification rate (CCR). The gold numbers are size = 2, TDR = 1, MCC = 1, and CCR = 1. Our NBI test tends to give slightly conservative results with smaller size, a known consequence of FDR control (*Benjamini and Hochberg*, 1995). Clearly, SAS and riskRFE pay a price of overfitting with a large number of noise features selected, resulting in larger size and smaller TDR. One lesson we learn from the simulation is that we may first run SAS or riskRFE to select a relatively large pool of potential biomarkers, and then apply NBI to control FDR. In this way, we could reach a desirable balance of sensitivity and specificity. In regard to MCC, the proposed NBI test outperforms SAS and riskRFE, except for the linear scenario, where SAS gives the highest MCC. In addition, the estimated ITR derived from the NBI method gives the highest CCR for the binary scenario, but not for the linear and nonlinear scenarios, which is a limitation of OWL. Some additional simulation experiments, including the small sample cases where $n = 200$, the scenarios where $\boldsymbol{X}_e = Null$ and etc. are included in Appendix D.

2.5 Analysis of Diabetes Trial Data

We apply the proposed NBI methodology to analyze the motivating diabetes trial described in Section 2.2. The outcome of benefit is the average reduction rate of

Table 2.3: Size, TDR, MCC, and CCR for variable selection based on NBI test, SAS and riskRFE in the multiple-variable-based decision rule evaluation.

NBI					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	800	1.745 (0.709)	0.906 (0.222)	0.801 (0.242)	0.835 (0.081)
	1000	1.813 (0.689)	0.922 (0.194)	0.828 (0.225)	0.852 (0.076)
	1200	1.904 (0.622)	0.934 (0.170)	0.868 (0.202)	0.870 (0.070)
binary	800	1.844 (0.747)	0.894 (0.232)	0.813 (0.248)	0.765 (0.098)
	1000	1.917 (0.705)	0.906 (0.215)	0.847 (0.232)	0.786 (0.095)
	1200	1.924 (0.650)	0.919 (0.199)	0.863 (0.222)	0.805 (0.090)
nonlinear	800	1.805 (0.744)	0.904 (0.220)	0.802 (0.245)	0.818 (0.081)
	1000	1.926 (0.738)	0.910 (0.199)	0.833 (0.225)	0.832 (0.075)
	1200	1.913 (0.616)	0.929 (0.186)	0.869 (0.217)	0.847 (0.073)
SAS					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	800	2.817 (0.887)	0.774 (0.212)	0.831 (0.165)	0.971 (0.012)
	1000	2.525 (0.732)	0.846 (0.193)	0.887 (0.146)	0.976 (0.010)
	1200	2.339 (0.609)	0.898 (0.169)	0.926 (0.126)	0.979 (0.010)
binary	800	3.662 (1.244)	0.613 (0.215)	0.693 (0.187)	0.743 (0.014)
	1000	3.286 (1.095)	0.677 (0.219)	0.751 (0.180)	0.744 (0.014)
	1200	3.052 (1.007)	0.723 (0.217)	0.790 (0.174)	0.744 (0.014)
nonlinear	800	3.238 (1.102)	0.688 (0.219)	0.760 (0.180)	0.943 (0.013)
	1000	2.915 (0.942)	0.752 (0.215)	0.813 (0.169)	0.948 (0.012)
	1200	2.654 (0.817)	0.814 (0.205)	0.862 (0.157)	0.949 (0.011)
riskRFE					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	800	3.091 (1.003)	0.660 (0.222)	0.704 (0.222)	0.852 (0.056)
	1000	2.815 (0.895)	0.732 (0.221)	0.773 (0.209)	0.869 (0.057)
	1200	2.586 (0.781)	0.797 (0.218)	0.828 (0.204)	0.883 (0.052)
binary	800	3.498 (1.135)	0.636 (0.211)	0.716 (0.178)	0.737 (0.123)
	1000	3.200 (0.990)	0.686 (0.209)	0.761 (0.169)	0.755 (0.124)
	1200	2.909 (0.847)	0.745 (0.206)	0.810 (0.159)	0.779 (0.117)
nonlinear	800	3.142 (1.083)	0.651 (0.235)	0.689 (0.243)	0.834 (0.061)
	1000	2.837 (1.009)	0.723 (0.230)	0.755 (0.219)	0.844 (0.059)
	1200	2.576 (0.823)	0.791 (0.224)	0.812 (0.215)	0.857 (0.055)

FPG over the 52 weeks of treatment. The base ITR is driven by three variables $\boldsymbol{X}_e = \{\text{b.FPG}, \text{age}, \text{BMI}\}$. Among those candidate biomarkers listed in Section 2.2, we want to select some important ones and evaluate their added values to improve the existing ITR.

We first performed a prescreening of these candidate biomarkers using SAS. Under the cut-off point 0.01 for the proportion of the incremental sequential advantage, SAS selects five variables potentially useful to update ITR, including baseline HbA1c, fasting insulin, AST, triglycerides and creatinine, denoted by $\boldsymbol{X}_u^{\text{SAS}}$. The resulting decision rule is

$$\begin{aligned}\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}}) \quad &= 0.13 - 0.14\text{b.FPG} - 0.02\text{age} - 0.17\text{BMI} - 0.05\text{HbA1c} \\ &\quad - 0.12\text{b.fasting insulin} - 0.13\text{AST} + 0.04\text{triglycerides} \\ &\quad + 0.18\text{creatinine},\end{aligned}$$

We would allocate a patient with Type 2 diabetes to take pioglitazone if $\widehat{f} > 0$ and to take gliclazide if $\widehat{f} < 0$. $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}})$ assigns 586 patients to pioglitazone and 244 patients to gliclazide. Following *Murphy et al.* (2001), we further calculate the estimated value function by $\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))B/P(A \mid \boldsymbol{X})]/\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))/P(A \mid \boldsymbol{X})]$, where $\mathbb{E}_n^*$ is the empirical average value. In order to make the comparison from the same baseline, the same method (e.g. SVM, which is the learning algorithm for both NBI and riskRFE) is used to calculate the estimated value function. $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}})$ gives an estimated value function of 0.049, meaning that the expected average FPG reduction rate would be 0.049 mmol/L/week over 52 weeks if $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}})$ were implemented for the whole population. The estimated value functions given by complete random allocation and $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e)$ are 0.048 and 0.052 respectively, indicating that $\boldsymbol{X}_u^{\text{SAS}}$ does not improve the existing decision rule as far as the estimated value function concerns. We then performed a biomarker screening using riskRFE, which selected three variables potentially useful to update the existing

ITR, including baseline fasting insulin, creatinine and GGT. The estimated decision rule is

$$\begin{aligned}\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}}) \quad &= -1.48 - 0.08\text{b.FPG} + 0.37\text{age} + 1.09\text{BMI} \\ &\quad + 0.45\text{b.fasting insulin} + 0.53\text{creatinine} + 0.97\text{GGT},\end{aligned}$$

$\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}})$ assigns 527 patients to take pioglitazone and 303 patients to take gliclazide. The estimated value function given by $\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}})$ is 0.051, even slightly smaller than that given by $\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e)$, which is 0.052, indicating that $\boldsymbol{X}_u^{\text{riskRFE}}$ is not actually improving the existing ITR.

A clinically relevant question to the above estimated decision rules is whether the selected candidate variables in $\boldsymbol{X}_u^{\text{SAS}}$ and $\boldsymbol{X}_u^{\text{riskRFE}}$ are really influential to ITR, or whether we may further reduce $\boldsymbol{X}_u^{\text{SAS}}$ and $\boldsymbol{X}_u^{\text{riskRFE}}$ to achieve a more cost-effective ITR. We then applied the proposed NBI test. It turns out that baseline fasting insulin is the only candidate variable selected by our NBI method. This gives a simpler decision rule,

$$\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{NBI}}) = -1.30 - 0.07\text{b.FPG} + 0.39\text{age} + 2.04\text{BMI} + 0.35\text{b.fasting insulin},$$

which assigns 481 patients to pioglitazone and 349 patients to gliclazide. The estimated value function of this OWL-updated treatment regime is 0.053, which is higher than that given by $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}})$ and $\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}})$, although $\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{NBI}})$ only uses a single biomarker, baseline fasting insulin, to improve ITR. The estimated value functions given by $\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e)$ is 0.052, indicating that the inclusion of $\boldsymbol{X}_u^{\text{NBI}}$ in the decision rule also improves the ITR with respect to the expected average FPG reduction rate.

Comparing the coefficients in the estimated decision rules, we notice that the signs of the coefficients for some common variables are different in $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}})$,

$\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}})$ and $\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{NBI}})$ (e.g. age has a negative coefficient in $\widehat{f}_{\text{SAS}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{SAS}})$ but positive coefficients in $\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}})$ and $\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{NBI}})$). It may due to the following reasons. First, loading coefficients are estimated by conditioning on other variables in the decision rule, and thus signs of these coefficients are possibly different with different sets of predictors (which are correlated) used in the construction of the decision functions. What matters the most is indeed the maximum treatment benefit, which is the primary objective of this learning procedure. Although the signs of the coefficients are not directly interpretable in this type of methodology, we would still see a great deal of concordance, in particular between $\widehat{f}_{\text{riskRFE}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{riskRFE}})$ and $\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{NBI}})$. This is because that NBI and riskRFE are both based on support vector machine (SVM), while SAS is based on regularized linear regression. Thus, the training procedures and estimating criteria are different between SAS and NBI/riskRFE. With the inclusion of the selected variables $\boldsymbol{X}_u^{\text{NBI}}$, $\widehat{f}_{\text{NBI}}(\boldsymbol{X}_e, \boldsymbol{X}_u^{\text{NBI}})$ improves the estimated value function by about 10% compared to complete randomness. In regard to the clinical impact of our results, we would think that the demonstrated improvement is clinically meaningful, especially for people on the border line of diabetes, i.e. the so-called prediabetes. It is known that approximately 88 million adults - more than 1 in 3 - have prediabetes in the US. Of those with prediabetes, more than 80% don't know they have it. The 10% change may help pre-diabetic people whose diagnostic values just cross the border line to be controlled at the normal level.

In addition to the expected reduction rate of FPG, we also compare the expected reduction rate of HbA1c, another outcome of interest in the trial. Preliminary analysis identifies no significant difference between the HbA1c reduction rates for patients receiving pioglitazone or gliclazide. The inclusion of $\boldsymbol{X}_u^{\text{NBI}}$, $\boldsymbol{X}_u^{\text{SAS}}$ and $\boldsymbol{X}_u^{\text{riskRFE}}$ do not improve the existing ITRs with all the existing and updated decisions rules giving the same estimated reduction of HbA1c, which is 0.031 mmol/L/week over the 52-

week treatment. But it is still slightly higher than the value of 0.029 given by complete random allocation A.

2.6 Concluding Remarks

In this article, we proposed a new biomarker assessment tool, termed as NBI, that enables to evaluate added values of biomarkers for improving existing ITRs. This new method can be used in both single and multiple-variable-based decision rule evaluations. Extensive simulation studies demonstrate that our method can correctly identify signal biomarkers under various scenarios with desirable performances in comparison to existing methods. Application of the proposed method to a real diabetes clinical trial reveals that baseline fasting insulin is an important biomarker that can significantly improve an existing ITR involving age, BMI, and baseline fasting FPG, for the allocation of pioglitazone or gliclazide to patients with Type 2 diabetes. It results significant clinical benefit of average reduction rate of FPG during the 52 weeks of treatment.

NBI is an analog to net reclassification improvement (NRI), a seminal index that has been extensively used to evaluate the usefulness of new markers for predicting risk of developing diseases (*Pencina et al.*, 2010). The proposed NBI is fundamentally different from NRI in the sense that NBI pertains to reclassification with respect to treatment group when class labels are not directly observed, rather based on outcome of treatment. *Pepe et al.* (2014) demonstrated that false-positive conclusions based on the NRI statistic were unacceptably high. However, our simulation studies have illustrated that the false discovery rate is well controlled using the NBI method. *Vickers and Pepe* (2014) pointed out that NRI weights reclassification (i.e. false positive and false negative) inappropriately, which may also be an underlying problem of NBI. However, with no information of true label available in the setting of NBI, appropriate weighting of false positive and false negative may be infeasible in prac-

tice. Decision curve analysis (DCA) (*Vickers and Elkin*, 2006) is another commonly used method for comparing multiple treatment decision rules to select the optimal one that maximizes the outcome of interest. The formulation of DCA relies on the calculation of a net benefit, which is the relative harm of a false-positive and a false-negative. Similarly, it is infeasible to apply DCA in our setting since we never know the underlying true labels for patients in a clinical trial in practice.

NBI is naturally applicable for non-linear decision rules due to the invocation of SVM, in which different kernels (e.g. Gaussian kernel) can be easily applied. We would like to clarify two points in the usage of kernels: (i) kernel is used exclusively to model the conditional distribution of R given $\boldsymbol{X}$, and (ii) the assumption of additive errors in the generation of X_{null} is imposed on the conditional distribution of $\boldsymbol{X}_u$ given $\boldsymbol{X}_e$. Thus, the additivity assumption does not influence the relationship (or the decision rule) between R and $\boldsymbol{X}$ characterized by kernel in the generation of the empirical null distribution. In order to generate the null distribution for NBI, a certain assumption on the influence of random errors on signals is inevitable. In this paper, we adopted the classical additive error assumption, which can be violated in the practice. In order to check the validity of the additive error assumption, we suggest first to run a residual diagnosis. If it indicates that the error is not additive, i.e. $X_u = f(\boldsymbol{X}_e, \epsilon)$, we can apply Taylor Expansion to the function $f(\boldsymbol{X}_e, \epsilon)$ and use a generalized additive model (GAM) to model X_u on $\boldsymbol{X}_e$. Then permutation test can be performed based on the new error term $\epsilon' = X_u - \widehat{\text{GAM}}(\boldsymbol{X}_e)$, or $\epsilon'' = \{X_u - \widehat{\text{GAM}}(\boldsymbol{X}_e)\}/\sqrt{\widehat{\theta}(\boldsymbol{X}_e)}$ if we assume the variance of ϵ' can be modeled by a function $\theta(\boldsymbol{X}_e)$.

One future direction of this study is to extend this methodology to multiple treatment settings since clinical trials sometimes have more than two treatments in practice. In addition, it may be desirable to extend the method to situations where clinical benefit outcomes are categorical or time-to-event. With increasing interest and research in dynamic treatment regimes, we may also extend the NBI test to settings

with multiple decision time points. Due to the number of replicates required by the proposed permutation test as well as the computation time needed for SVM, the proposed method may run into high computational demand. The algorithm can become faster if a theoretically justified null distribution is available for the NBI test statistic, which is worth an exploration in our future research.

CHAPTER III

Synergistic Self-Learning Approach to Establishing Individualized Treatment Rules from Multiple Benefit Outcomes

3.1 Introduction

According to the developmental origins of health and disease (DOHaD) hypothesis, environmental exposure before and immediately after birth may influence the developmental health and well-beings of children during their infancy and even later in their adulthood (*Wadhwa et al.*, 2009). Prenatal exposure to toxic agents, maternal lifestyle, and maternal nutrition intake have been reported as major factors that directly affect embryonic development. A vast literature has unveiled that lead is detrimental on human neurobehavioral and cognitive development, particularly on children (*Chen et al.*, 2005; *Hornung et al.*, 2009; *Wasserman et al.*, 2003; *Téllez-Rojo et al.*, 2006b; *Hu et al.*, 2006; *Braun et al.*, 2012). It is demonstrated that both blood and bone lead levels of children are associated, in an inverse direction, with their intelligence. Therefore, it is important to control lead exposure to children in order to protect their neurodevelopment. In addition to the environmental lead exposure after birth, the *in utero* exposure to lead that fetuses receive comes solely from maternal nutrients. Thus, the study of maternal lead control during their pregnancy is of great

importance for maternal and child health.

One of the leading cohort studies in this field is the Early Life Exposure in Mexico to ENvironmental Toxicants (ELEMENT) Project (*Perng et al.*, 2019). The ELEMENT project is an over 25-year mother–child pregnancy and birth cohort carried out in Mexico, where regulations to prevent the use of lead in commodities such as pigments, glaze, paint as well as gasoline were established only in 1991, resulting in a serious problem of lead exposure to pregnant women even in late 20th Century compared to the USA, where the control of lead usage started since 1970. One of the major objective of ELEMENT is to study the effect of maternal lead exposure on infant health outcomes. A lot of previous studies have discovered that mobilization of maternal cumulative lead storage from bones (e.g. patella and tibia) into blood circulation (*Gulson et al.*, 1995, 2003; *Hu and Hernandez-Avila*, 2002) is one of the important sources that contribute to maternal blood lead levels. Thus, blocking or reducing the amount of lead released into maternal blood circulation from bones is a significant preventive measure to minimize fetal lead exposure, and this may be achieved by calcium, a lead blocker. Many studies, including several published works from the ELEMENT Project, have demonstrated that dietary calcium supplementation can reduce maternal lead concentrations in blood by reducing bone resorption of women during pregnancy (*Ettinger et al.*, 2005, 2007; *Janakiraman et al.*, 2003; *Gulson et al.*, 2004; *Téllez-Rojo et al.*, 2006a). All these studies focus on the population-wise effect of calcium supplementation on reducing the blood lead concentration of pregnant women, which ignores the heterogeneity of individual response to calcium supplementation. Clearly, the calcium treatment effect size varies from one mother to another; one taking the calcium supplement may have a significant reduction in blood lead, while another in the same treatment arm may not have a strong effect due to various reasons, e.g. there is enough daily dietary calcium intake. Thus, in addition to evaluating the population-average effect of calcium on the whole study

participants, establishing an individualized treatment rule (ITR) is of clinical importance to pregnant women, useful in precision nutritional counselling and guidelines on tailored individual dietary patterns, lifestyle and living conditions.

The major objective of this article is to establish an ITR for pregnant women based on their characteristic features to take calcium supplementation in order to maximize the reduction of maternal lead exposure to infants. The establishment of a well-performing and robust ITR requires a large sample size of high quality data with measurements of the clinical outcome we want to maximize, i.e. the reduction of lead exposure to infants in the ELEMENT Project. As we can see from the ELEMENT data (more details will be discussed in Section 3.2), there is only a small number of measurements of the primary clinical outcome, the cord blood lead concentration, due to biosample collection difficulties. The derived ITR based on such a small amount of data will definitely suffer from its robustness and reproducibility. A straightforward idea to solve this problem is to integrate some secondary outcomes of interest, such as maternal blood lead concentration from the third trimester and infant blood lead concentration at 3-month age, with large sample sizes into the estimation of ITR. In a randomized clinical trial, it is not uncommon that in addition to the primary outcome, many secondary outcomes such as quality-of-life measurements are typically collected for the evaluation of broad treatment benefits. With multiple benefit outcomes measured, two major challenges arise in ITR derivation. On one hand, primary and secondary outcomes are of different clinical relevance to the overall underlying health benefit. On the other hand, outcomes may be measured separately from different subgroups of subjects, leading to different sample sizes and complicated missing data patterns. Existing decision rule estimation methods in the current literature, such as Q-learning (*Qian and Murphy*, 2011), outcome weighted learning (OWL) (*Zhao et al.*, 2012), and residual weighted learning (RWL) (*Zhou et al.*, 2017) are applicable only for a single outcome related to health benefit. To our

best knowledge, there are no existing methods available in the literature that permits the integration of multiple outcomes of different clinical relevance and data missingness into a decision rule derivation. In order to fill in this important methodological gap and address this technical need, we propose a new method, named synergistic self-learning (SS-learning), that can incorporate both multiple outcomes and incomplete data in the derivation of ITR.

The SS-learning method works in a weighted self-learning paradigm with characterization of the relative contributions of respectively low- and high-quality data sources. We analyzed the calcium supplementation trial from the ELEMENT Project using SS-learning to derive an ITR for calcium supplementation assignment to pregnant women. Results based on SS-learning illustrate that 3-month infant blood lead concentration is only slightly down weighted, while maternal blood lead concentration at the third trimester has a much lower clinical relevance, compared with cord blood level in the derivation of ITR. The resulted ITR gives the highest estimated lead exposure reduction compared with standard OWL either on the cord blood lead outcome alone or on the combined outcomes with naive equal weighting (i.e. a naive assumption of their equal clinical relevance). Our findings indicate that SS-learning produces the largest reduction of infant blood lead concentration should the estimated ITR be implemented to the whole study population. This optimality achieved through personalized calcium supplementation allocation is clinically appealing in precision nutrition. Moreover, we also identified several variables playing significant roles in calcium supplement allocation, including dietary fiber, Fe, Zn and vitamin C intake, indicating complicated interactions of nutrient contents in reducing maternal lead exposure to infants. Although the focus of this article is in ITR derivation for the calcium supplementation trial from the ELEMENT project, the developed SS-learning method also contributes to the general statistical literature. Extensive simulations show that it is very effective to improve classification rate with the use of

the augmented training data. In addition, the algorithmic convergence of SS-learning has also been proved.

The remainder of the chapter is organized as follows. We present an overview of the calcium supplementation trial from the ELEMENT Project in Section 3.2. We introduce SS-learning in Section 3.3. In Section 3.4 we evaluated the SS-learning method through simulation studies with comparisons to existing methods, followed by Section 3.5 where we analyzed data of the ELEMENT calcium trial with interpretations and discussions. We give some concluding remarks in Section 3.6. Additional numerical results and proof of the algorithmic convergence of SS-learning can be found in Appendix E-G.

3.2 Application: Calcium Supplementation Trial

The data we analyzed is generated from the clinical trial in the third cohort study of the ELEMENT Project. This clinical trial aims to assess effects of daily calcium supplementation for pregnant women on reducing fetal blood lead concentration. A total of 670 women recruited at their first trimester during 2001 to 2003 in Mexico City participated in the clinical trial, with 334 participants randomized to receive a daily dose of 1200 mg calcium supplement and 336 randomized to receive a placebo. The maternal baseline variables used to establish the ITR include age, weight, total years in school (school), status of marriage (married: yes/no), hrp parity (parity), dietary fiber, Fe, Zn and vitamin C (VC) intake, and hemoglobin concentration (HgB). Blood lead concentration measured from mother at first trimester (PBM) is also included as a baseline reference of the maternal lead level. Maternal dietary calcium intake (Ca) is included as a confounding factor for calcium supplementation.

Various benefit outcomes were measured during or after the trial, including maternal blood lead concentration at different trimesters of pregnancy, cord blood lead concentration, and infant blood lead concentration up to 3-month of age. The primary

measurement of fetal lead exposure is cord blood lead concentration since it directly reflects the dose of lead in the exchange of nutrients between fetus and mother. Unfortunately, due to logistic difficulties of biosample collection, only a small number of participants ($n_1 = 60$) contributed cord blood samples at birth, leading to a limited number of cord blood lead measurements. Deriving an ITR based on such a small training dataset is at high stakes of reproducibility. One way to increase the size of training data is to embrace other types of lead exposure measurements, albeit possibly of lower clinical relevance. This motivates us to consider blood lead concentration measured from infants at 3-month age (PBC3), as well as maternal blood lead concentration measured at the third trimester (PBM3), both of which are clinically much easier to capture, giving rise to a much larger sample size ($n_2 = 245$ and $n_3 = 366$). In comparison to the cord blood lead measurement, blood lead concentration at age 3 months may be slightly "contaminated". Biologically, the effect of fetal lead exposure from mother diminishes slowly overtime and may be altered by other sources of lead in environment and breast milk; see Figure 3.1 for the mechanistic pathway of fetal exposure to lead. Meanwhile, maternal blood lead concentration at the third trimester do not directly reflect the lead exchange between mother and fetus during pregnancy. Therefore, infant's 3-month blood lead concentration and maternal blood lead concentration at the third trimester are less relevant outcomes compared to cord blood lead, as far as the efficacy of calcium supplementation concerns. In order to derive an optimal ITR with improved reliability and reproducibility, we need a training dataset not only of high relevance and precision to clinical outcome but also of large sample size. To do so, we would like to apply an ITR estimation methods that accommodate multiple outcomes of different relevance and sample sizes. The summary statistics for the predictors used to establish the ITR are listed in Table 3.1.

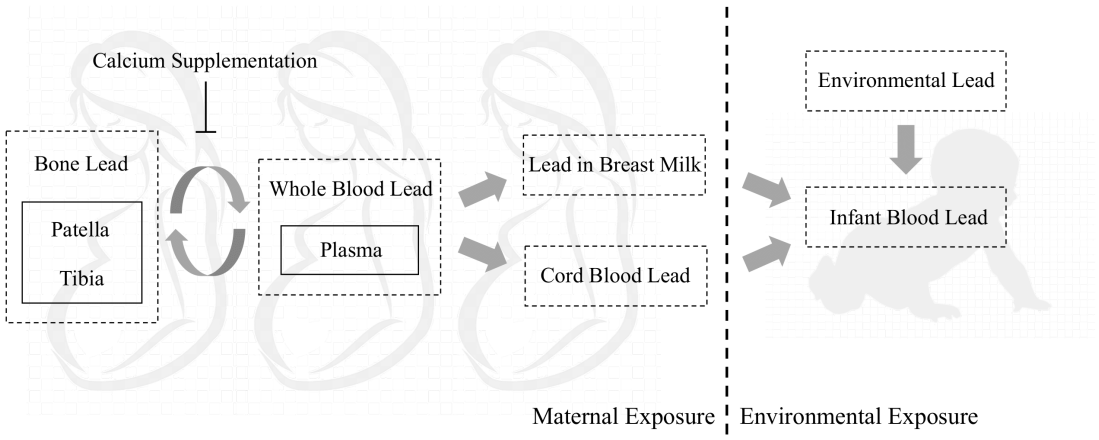

Figure 3.1: Mechanistic pathway of fetal exposure to lead. During pregnancy, the increasing maternal bone turnover results in an elevated lead release into plasma, which causes the increasing lead concentrations in cord blood and breast milk. Such maternal lead exposure together with environmental lead exposure result in an overall detrimental effect on the neurobehavioral and cognitive development of infants.

Table 3.1: Summary statistics of the predictors in the calcium supplementation trial. Mean (sd) and percentage values are shown, where p-values are obtained from t-test and chi-square test for the numeric and categorical variables, respectively.

Predictor	Calcium	Placebo	p-value
age	26.9 (5.6)	25.9 (5.3)	0.017
weight	62.0 (10.7)	61.7 (10.0)	0.672
school	10.8 (2.9)	10.6 (2.9)	0.440
married	0.689	0.685	0.975
parity	2.0 (1.0)	2.1 (1.1)	0.409
HgB	13.2 (1.1)	13.2 (1.1)	0.774
Ca	11.1 (4.9)	10.8 (5.3)	0.531
fiber	24.2 (9.3)	23.4 (9.0)	0.237
Fe	13.4 (5.5)	13.4 (5.6)	0.995
Zn	9.7 (3.5)	9.7 (3.6)	0.893
VC	182.8 (90.4)	177.4 (86.5)	0.426
PBM	4.6 (2.6)	5.2 (3.7)	0.065

3.3 Formulation of SS-learning

Here we first introduce the notations and basic settings of ITR derivation followed by an introduction of a widely used decision rule estimation method OWL. Then we propose the SS-learning method that can be applied to ITR estimation with multiple outcomes with the parameter tuning procedure and the algorithmic convergence theorem.

3.3.1 Basic Setting

Consider a training dataset $\mathcal{S} = \{(B_{1i}, B_{2i}, \ldots, B_{mi}, A_i, \boldsymbol{X}_i), i = 1, \ldots, n\}$ collected from a two-armed randomized clinical trial. $\boldsymbol{X} = (X_1, \ldots, X_d)^\intercal \in \mathcal{X} \subseteq \mathbb{R}^d$ is a d-dimensional vector of features. $A \in \mathcal{A} = \{-1, +1\}$ is a treatment (e.g. calcium supplementation or placebo) assigned to each subject with probability $P(A \mid \boldsymbol{X})$. With complete randomization, $P(A = 1 \mid \boldsymbol{X}) = P(A = -1 \mid \boldsymbol{X}) = 0.5$. But the propensity score $P(A \mid \boldsymbol{X})$ can also be estimated by methods such as logistic regression. $(B_1, B_2, \ldots, B_m)$ are multiple benefit outcomes. We assume that B_1 is of high relevance but difficult to measure, while $(B_2, \ldots, B_m)$ are of low relevance but easy to measure. Let $\mathcal{S}_1 = \{(B_{1i}, A_i, \boldsymbol{X}_i), i = 1, \ldots, n_1\}$ be the subset of participants with B_1 measured, regardless of $(B_2, \ldots, B_m)$ being measured or not. Let $\mathcal{S}_l = \{(B_{li}, A_i, \boldsymbol{X}_i), i = 1, \ldots, n_l\}, l = 2, \ldots, m$ be the subsets of participants with only B_l measured. With no surprise, $n_1 = |\mathcal{S}_1|$ is much smaller than $n_l = |\mathcal{S}_l|, l = 2, \ldots, m$ since it is more difficult to obtain measurements of B_1. Our objective is to estimate the optimal ITR that assigns treatment $A \in \{-1, +1\}$ to each subject based on personal features $\boldsymbol{X}$ with the aim of maximizing the expected benefits.

3.3.2 Synergistic Self-learning

In the case of a single benefit outcome, *say* B, the existing seminal OWL (*Zhao et al.*, 2012) estimates an optimal decision rule D^* by solving the following optimiza-

tion function (see details of OWL algorithm in Appendix A and B):

$$\begin{aligned} \min_{\boldsymbol{\omega},b,\boldsymbol{\xi}} \quad & \frac{1}{2}||\boldsymbol{\omega}||^2 + C\sum_{i=1}^{n} \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)}\xi_i \\ \text{s.t.} \quad & A_i(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \xi_i \,\text{and}\, \xi_i \geq 0,\ i = 1, ..., n \end{aligned} \tag{3.1}$$

where $f(\boldsymbol{X}) = \boldsymbol{\omega}^\intercal \boldsymbol{X} + b$ is the decision rule over the space of linear functions, C is the tuning parameter of SVM, and $\xi_i = (1 - A_i f(\boldsymbol{X}_i))^+$ is the slack variable. From (3.1) we see that the main difference between OWL and standard SVM is that the former weights each slack variable ξ_i with a scaled personal clinical benefit value $B_i/P(A_i \mid \boldsymbol{X}_i)$. We can generalize the approach to obtain a nonlinear decision rule using a kernel function $\mathcal{K} : \mathcal{X} \times \mathcal{X} \to R$ via the theorem of the *reproducing kernel Hilbert space* (RKHS) $\mathcal{H}_\mathcal{K}$ (*Cortes and Vapnik*, 1995). For simplicity, we focus on linear decision rules in the following algorithm. But the method can be easily extended to nonlinear decision rules with little effort.

OWL only works for a single outcome, which has to be extended to analyze data with multiple outcomes. To address this technical need, we propose the following algorithm.

Algorithm III.1. Suppose that we have one high-quality dataset $\mathcal{S}_1 = \{(B_{1i}, A_i, \boldsymbol{X}_i), i = 1, \ldots, n_1\}$ and multiple low-quality datasets $\mathcal{S}_l = \{(B_{li}, A_i, \boldsymbol{X}_i), i = 1, \ldots, n_l\}$, $l \in \mathcal{L} = \{2, \ldots, m\}$, where $n_1 < n_l$ and B_1 is of higher clinical relevance than B_l. Let $p = P(A \mid \boldsymbol{X})$. We iterate the following steps. Refer to Figure 3.2 for the workflow of the proposed SS-learning algorithm.

S1 Estimate an ITR by OWL using $\mathcal{S}_1$ of high quality and get the initial estimates of the parameters $(\boldsymbol{\omega}^{(0)}, b^{(0)}, \boldsymbol{\xi}^{(0)})$ from (3.1). Then predict labels for individuals in $\bigcup_{l=2}^{m} \mathcal{S}_l$ of low qualities. Denote the predicted labels as $y_{li}^{(0)}, i \in \mathcal{S}_l, l \in \mathcal{L}$.

S2 The k-th ($k \geq 1$) iteration runs through the following Steps 2.1-2.3.

S2.1 Define an augmented training dataset $\mathcal{S}^{(k)} = \mathcal{S}_1 \cup \tilde{\mathcal{S}}_2^{(k)} \cup \cdots \cup \tilde{\mathcal{S}}_m^{(k)}$, where individuals in $\tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}$ have their predicted labels from the previous iteration $k-1$, resulting in $\tilde{\mathcal{S}}_l^{(k)} = \{(B_{li}, A_i, y_{li}^{(k-1)}, \boldsymbol{X}_i), i = 1, \ldots, n_l\} = \mathcal{S}_l \cup \{y_{li}^{(k-1)}, i = 1, \ldots, n_l\}, l \in \mathcal{L}$.

S2.2 Solve the following optimization problem using $\mathcal{S}^{(k)}$:

$$\begin{aligned} \min_{\boldsymbol{\omega}, b, \boldsymbol{\xi}} \quad & \frac{1}{2}||\boldsymbol{\omega}||^2 + C\left[\sum_{i \in \mathcal{S}_1} \frac{B_{1i}}{p_i}\xi_i \right. \\ & \left. + \sum_{l=2}^{m}\left\{\lambda_l \sum_{i \in \tilde{\mathcal{S}}_l^{(k)}} \frac{B_{li}}{p_i}\widetilde{\xi}_{li} + (1-\lambda_l)\sum_{i \in \tilde{\mathcal{S}}_l^{(k)}} \tau_{B_l}\check{\xi}_{li}\right\}\right] \\ \text{s.t.} \quad & A_i(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0,\ i \in \mathcal{S}_1, \\ & A_i(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \widetilde{\xi}_{li} \text{ and } \widetilde{\xi}_{li} \geq 0,\ i \in \tilde{\mathcal{S}}_l^{(k)},\ l \in \mathcal{L}, \\ & y_{li}^{(k-1)}(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \check{\xi}_{li} \text{ and } \check{\xi}_{li} \geq 0,\ i \in \tilde{\mathcal{S}}_l^{(k)},\ l \in \mathcal{L}. \end{aligned} \tag{3.2}$$

where $\lambda_l \in [0, 1]$ is an additional tuning parameter that characterizes the relative contribution of the predicted and observed information from $\tilde{\mathcal{S}}_l^{(k)}$ in problem (3.2). τ_{B_l} is a summary statistic (e.g. sample mean) of $\frac{B_l}{p}$ that is added as the weight for $\check{\xi}_{li}$ to balance the magnitude between $\sum_{i \in \tilde{\mathcal{S}}_l^{(k)}} \widetilde{\xi}_{li}$ and $\sum_{i \in \tilde{\mathcal{S}}_l^{(k)}} \check{\xi}_{li}$, so to reduce the bias over the tuning of λ_l. Denote the estimates of the parameters as $(\boldsymbol{\omega}^{(k)}, b^{(k)}, \boldsymbol{\xi}^{(k)})$ at iteration k. Also, the predicted labels for $\tilde{\mathcal{S}}_l^{(k)}$ are updated as $y_{li}^{(k)}, i \in \tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}$.

S2.3 Calculate the objective function at the estimates $(\boldsymbol{\omega}^{(k)}, b^{(k)}, \boldsymbol{\xi}^{(k)})$ as

$$\begin{aligned} h(\boldsymbol{\omega}^{(k)}, \boldsymbol{\xi}^{(k)}) = & \frac{1}{2}||\boldsymbol{\omega}^{(k)}||^2 + C\left[\sum_{i \in \mathcal{S}_1} \frac{B_{1i}}{p_i}\xi_i^{(k)} \right. \\ & \left. + \sum_{l=2}^{m}\left\{\lambda_l \sum_{i \in \tilde{\mathcal{S}}_l^{(k)}} \frac{B_{li}}{p_i}\widetilde{\xi}_{li}^{(k)} + (1-\lambda_l)\sum_{i \in \tilde{\mathcal{S}}_l^{(k)}} \frac{\overline{B_l}}{p}\check{\xi}_{li}^{(k)}\right\}\right] \end{aligned} \tag{3.3}$$

S3 The algorithm stops if $\frac{|h(\boldsymbol{\omega}^{(k)},\boldsymbol{\xi}^{(k)})-h(\boldsymbol{\omega}^{(k-1)},\boldsymbol{\xi}^{(k-1)})|}{h(\boldsymbol{\omega}^{(k-1)},\boldsymbol{\xi}^{(k-1)})} < \epsilon$ for a pre-determined precision constant ϵ, *say* 10^{-4}. The convergence values of $(\boldsymbol{\omega}^{(k)}, b^{(k)})$ are denoted by $(\hat{\boldsymbol{\omega}}, \hat{b})$, and the predicted labels at convergence for $\mathcal{S}$ are denoted as $\hat{y}_i, i = 1, \ldots, n$.

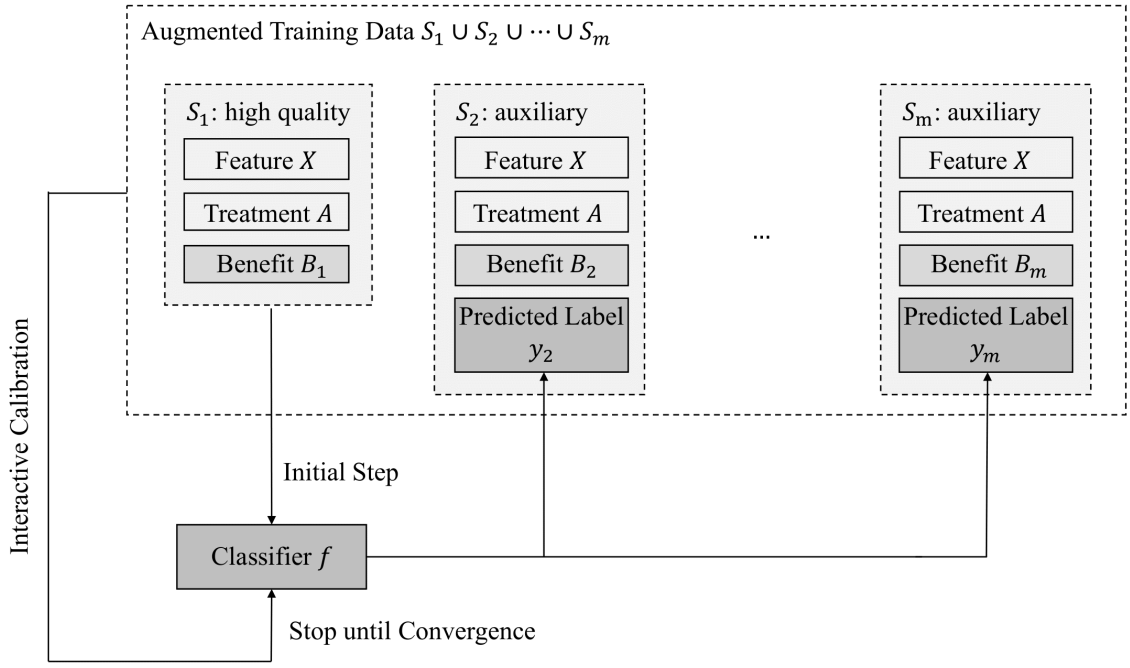

Figure 3.2: Workflow of the proposed synergistic self-learning algorithm.

The optimization function (3.2) is composed of three parts: (i) $\sum_{i\in\mathcal{S}_1} \frac{B_{1i}}{P(A_i|\boldsymbol{X}_i)}\xi_i$ represents an OWL problem based on $\mathcal{S}_1$; (ii) $\sum_{i\in\tilde{\mathcal{S}}_l^{(k)}} \frac{B_{li}}{P(A_i|\boldsymbol{X}_i)}\widetilde{\xi}_{li}$ represents an OWL problem based on $\tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}$; and (iii) $\sum_{i\in\tilde{\mathcal{S}}_l^{(k)}} \overline{\frac{B_l}{P(A|\boldsymbol{X})}}\check{\xi}_{li}$ represents a standard SVM classification problem based on $\tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}$ with the predicted labels $y^{(k-1)}$. During every iteration, each case i in $\tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}$ is associated with two pieces of information: a) benefit B_{li} resulted from treatment A_i, and b) predicted label $y_i^{(k-1)}$ that borrows information from $\mathcal{S}_1$. The optimization problem (3.2) takes both into consideration, where tuning parameter $\lambda_l \in [0, 1], l \in \mathcal{L}$ governs their relative importance. When $\lambda_l = 0$, we ignore the contribution from observed benefit B_l, and think that it is com-

pletely untrustworthy, and use the predicted labels only in optimization. In contrast, when $\lambda_l = 1$, we believe that B_l has the same quality as B_1, so to ignore the predicted labels, and consequently treat the problem (3.2) as a standard OWL on the combined data $\mathcal{S}_1 \cup \mathcal{S}_2 \cup \cdots \cup \mathcal{S}_m$.

3.3.3 Algorithmic Convergence

We establish the algorithmic convergence of SS-learning in the following theorem with the proof presented in Appendix E.

Theorem III.1. *The SS-learning objective function $h(\boldsymbol{\omega}^{(k)}, \boldsymbol{\xi}^{(k)})$ at iteration k satisfies the descending property over iterations, namely $h(\boldsymbol{\omega}^{(k)}, \boldsymbol{\xi}^{(k)}) \leq h(\boldsymbol{\omega}^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$ for $k \geq 2$. The equality occurs when the algorithm converges.*

3.3.4 Tuning Parameter Selection

To perform SS-learning, we need to choose the tuning parameters C and $\boldsymbol{\lambda}$. Due to a small sample size of the "good" data $\mathcal{S}_1$, the method of cross-validation does not yield stable and reliable tuning results. We propose to choose the values of C and $\boldsymbol{\lambda}$ at which the SSE of the predicted benefit values is minimized. This may be done by performing a grid search of C and $\boldsymbol{\lambda}$ on the entire dataset $\mathcal{S}_1 \cup \cdots \cup \mathcal{S}_m$. We define the SSE of the predicted benefit values as

$$\text{SSE} = \sum_{l=1}^{m} \left\{ \sum_{i=1}^{n_l} (B_{li} - \hat{B}_{li})^2 \right\}. \tag{3.4}$$

The individual predicted benefit value $\hat{B}_{li}$ is obtained from the following regression models,

$$B_l = \phi_l(\boldsymbol{X}, A\hat{f}(\boldsymbol{X})) + \epsilon_l, \tag{3.5}$$

where $\phi_l(a, b) = g_l(a) + \beta_l b$ for some function $g_l(\cdot), l = 1, \ldots, m$, $\hat{f}(\cdot)$ is the estimated decision function as an output of SS-learning, and $\epsilon_l, l = 1, \ldots, m$ are error terms.

In the simulation experiment described in Section 3.4 and the real data analysis discussed in Section 3.5, we choose linear models for ϕ_l to predict benefit. More general models like generalized additive model (GAM) can also be applied for ϕ_l (see the simulation experiment described in Appendix G) when the association between $\boldsymbol{X}$ and the benefit is assumed to be complicated and nonlinear.

3.4 Simulation Experiments

We ran multiple simulation experiments to validate the use of SS-learning in our data analyses. Here we report one of such simulation experiments to illustrate the finite sample performance of SS-learning in ITR derivation. More simulation results are included in Appendix F-G. We designed the simulation with three data sources: $\mathcal{S}_1$ contains high-quality benefit B_1, $\mathcal{S}_2$ and $\mathcal{S}_3$ contain low-quality benefit B_2 and B_3, respectively.

We randomly generated a 10-dimensional feature vector $\boldsymbol{X} = (X_1, \ldots, X_{10})^\intercal \in \mathbb{R}^{10}$ for each subject, where $X_\nu \overset{i.i.d}{\sim} U(0,1), \nu = 1, \ldots, 10$. Treatment $A \in \{-1, 1\}$ was randomly allocated to subjects with $P(A = 1 \mid \boldsymbol{X}) = P(A = -1 \mid \boldsymbol{X}) = 0.5$. The underlying true decision function f was specified as $f(\boldsymbol{X}) = 1 + X_1 + X_2 - 1.8X_3 - 2.2X_4$. The observed benefit B_1 was generated from a normal distribution with mean $\mu_1 = 0.01 + 0.02X_4 + 3.0Af(\boldsymbol{X})$ and standard deviation $\sigma_1 = 0.1$, the observed benefit B_2 was generated from a normal distribution with mean $\mu_2 = 0.1 + 0.2X_4 + 1.0Af(\boldsymbol{X})$ and standard deviation $\sigma_2 = 0.5$, while B_3 was generated from a normal distribution with mean $\mu_3 = 0.1 + 0.2X_4 + \beta_3 Af(\boldsymbol{X})$ and standard deviation σ_3. Let the total sample size $n = n_1 + n_2 + n_3$, where $n_1 = |\mathcal{S}_1|, n_2 = |\mathcal{S}_2|$ and $n_3 = |\mathcal{S}_3|$. In order to evaluate the performance of SS-learning under varying relevance of B_1, B_2 and B_3, we considered the following scenarios: (a) $\beta_3 \in \{0.1, 0.5, 1.0, 3.0\}, \sigma_3 = 1.0, n_1 = 20, n_2 = 100, n_3 = 380$; (b) $\beta_3 = 0.5, \sigma_3 \in \{0.1, 0.5, 1.0, 2.0\}, n_1 = 20, n_2 = 100, n_3 = 380$; (c) $\beta_3 = 0.5, \sigma_3 = 1.0, n_1 \in \{20, 100, 180, 250\}, n_2 = 100, n_3 \in \{380, 300, 220, 150\}$. In

addition, an independent dataset of sample size 1000 with true labels by $f(\boldsymbol{X}) > 0$ or otherwise was generated for external validation. The optimization function was solved using the linear kernel. We set tuning parameter $C \in \{1/128, 1/64, \ldots, 1/2, 1\}$ and weighting parameter $\lambda \in \{0, 0.1, \ldots, 0.9, 1.0\}$ for parameter selection. Simple linear regression with all the main effects of $\boldsymbol{X}$ was employed for the regression models $g_l(\boldsymbol{X}), l = 1, 2, 3$ in parameter tuning. For each scenario, simulation was repeated for 200 times to draw summary statistics.

We compared four methods for their prediction accuracy on the training and validation datasets: (M1) SS-learning with tuning according to the minimization of SSE; (M2) oracle SS-learning with optimal tuning according to the highest prediction accuracy on $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$; (M3) OWL with 5-fold cross-validation based tuning of C in SVM on $\mathcal{S}_1$ alone; (M4) OWL with 5-fold cross-validation based tuning of C in SVM on $\mathcal{S}$. M2 is regarded as the oracle method; the tuning of C and $\boldsymbol{\lambda}$ through the prediction accuracy of $\mathcal{S}$ is infeasible in reality because we never know the true underlying labels in $\mathcal{S}$. Both M3 and M4 are based on 5-fold cross-validation proposed in *Zhao et al.* (2012). Figure 3.3 shows the average prediction accuracy of these four methods on the training and validation datasets. It is shown that the oracle SS-learning M2 always gives the highest prediction accuracy. The M1 method, which we can use in practice, gives a slightly lower prediction accuracy compared to M2. The prediction accuracy given by M4 is constantly lower than M1 and M2. M3 always obtains the smallest prediction accuracy.

In addition, Table 3.2 lists the $\boldsymbol{\lambda}$ values tuned by method (3.4). Generally speaking, the selected λ_3 value increases with increasing β_3 and decreases with increasing σ_3. This is fully expected since larger β_3 and/or smaller σ_3 lead to higher relevance and better quality of B_3. However, the change of λ_3 along with different sample size ratios is not clearly shown due to the complicated relative clinical relevance between B_1, B_2 and B_3 represented by the relative sample sizes. The selected values

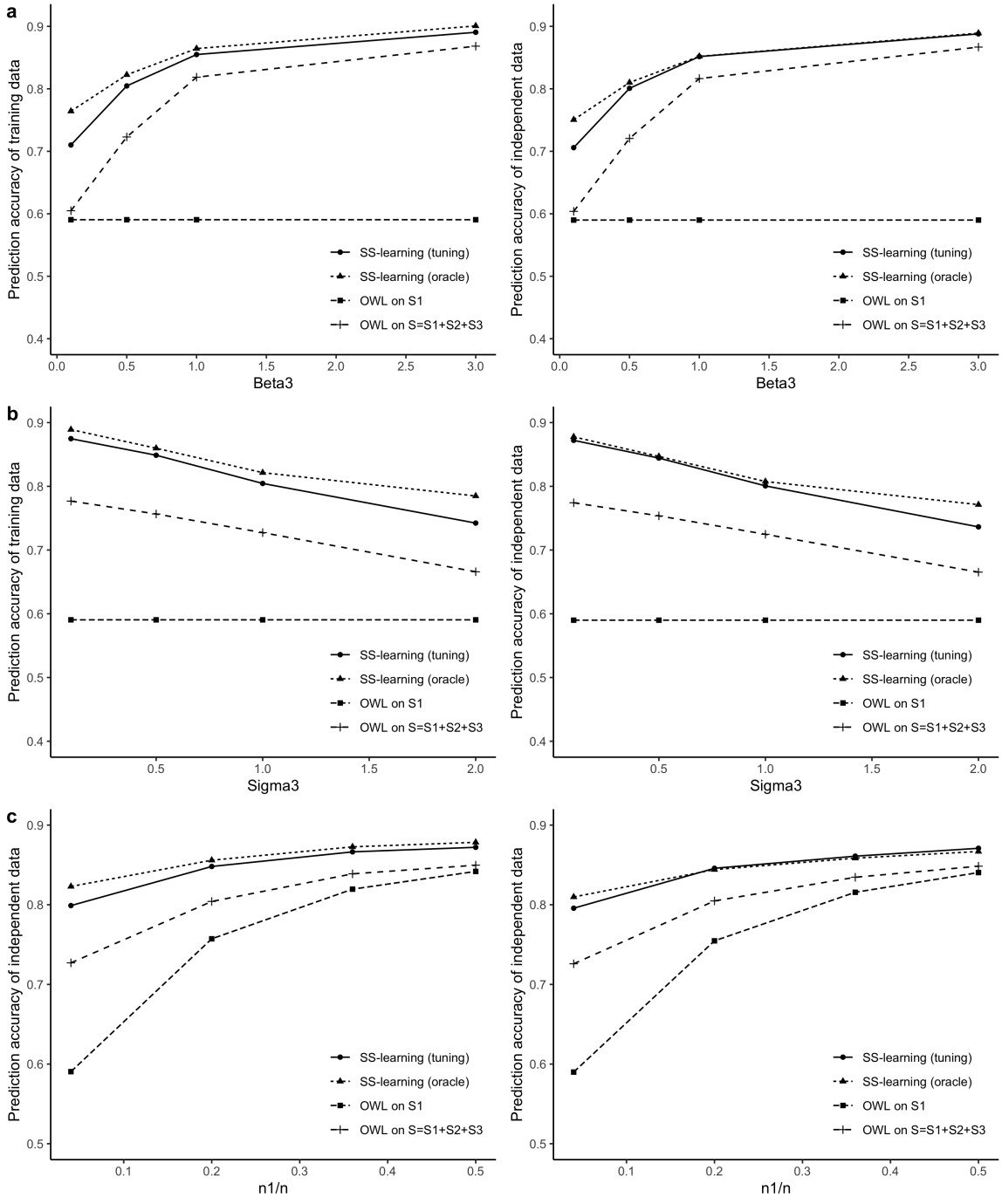

Figure 3.3: Comparison of prediction accuracy on the training and validation datasets by different methods under (a) varying β_3, (b) varying σ_3, (c) varying n_1/n.

Table 3.2: Selected weighting values λ_2 and λ_3 (mean (sd)) based on SS-learning using proposed tuning.

scenario	λ_2	λ_3
$\beta_3 = 0.1$	0.50 (0.37)	0.40 (0.32)
$\beta_3 = 0.5$	0.55 (0.38)	0.70 (0.28)
$\beta_3 = 1.0$	0.52 (0.37)	0.80 (0.23)
$\beta_3 = 3.0$	0.51 (0.38)	0.79 (0.22)
$\sigma_3 = 0.1$	0.45 (0.36)	0.90 (0.17)
$\sigma_3 = 0.5$	0.53 (0.38)	0.81 (0.22)
$\sigma_3 = 1.0$	0.55 (0.37)	0.68 (0.30)
$\sigma_3 = 2.0$	0.53 (0.37)	0.51 (0.30)
$n_1/n = 0.04$	0.56 (0.38)	0.66 (0.29)
$n_1/n = 0.20$	0.61 (0.37)	0.64 (0.34)
$n_1/n = 0.36$	0.51 (0.39)	0.64 (0.37)
$n_1/n = 0.50$	0.58 (0.41)	0.64 (0.38)

of λ_2 are stable over the changes of β_3, σ_3 and n_1/n, which is fully expected since the change of B_1 and B_3 barely changes the relative importance of B_2. The demonstrated sensitivity to varying clinical relevance and quality of auxiliary data signifies the proposed tuning method (3.4) as a satisfactory and reliable tuning procedure. We then calculated the estimated value functions of B_1, B_2 and B_3. The values are listed in Table 3.3-3.5. It is shown that M2 always gives the highest estimated value function as expected. But M1 gives higher estimated value functions than the OWL methods M3 and M4, indicating that the estimated ITR derived from M1 will lead to the highest clinical benefit if it were implemented for the whole study population in the future. Additional simulation results about the algorithmic convergence, the SSE values and the computation time based on methods M1-M4, as well as another nonlinear simulation experiment, are included in Appendix F-G.

3.5 Analysis of Calcium Supplementation Trial

In this section, we present the data analysis for the ELEMENT calcium supplementation trial using the SS-learning method. As pointed in Section 3.1, the central

Table 3.3: Estimated value function of B_1 (mean (sd)) based on SS-learning with proposed tuning, SS-learning with oracle tuning, OWL on $\mathcal{S}_1$, and OWL on $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$.

	SS-learning (tune)	SS-learning (oracle)	OWL on $\mathcal{S}_1$	OWL on $\mathcal{S}$
$\beta_3 = 0.1$	6.71 (1.34)	6.79 (1.26)	6.12 (1.23)	6.13 (1.47)
$\beta_3 = 0.5$	7.06 (1.28)	7.04 (1.29)	6.12 (1.23)	6.75 (1.37)
$\beta_3 = 1.0$	7.28 (1.31)	7.28 (1.26)	6.12 (1.23)	7.18 (1.28)
$\beta_3 = 3.0$	7.38 (1.38)	7.42 (1.31)	6.12 (1.23)	7.36 (1.28)
$\sigma_3 = 0.1$	7.22 (1.31)	7.24 (1.22)	6.12 (1.23)	6.90 (1.35)
$\sigma_3 = 0.5$	7.17 (1.23)	7.19 (1.24)	6.12 (1.23)	6.85 (1.35)
$\sigma_3 = 1.0$	7.05 (1.26)	7.05 (1.28)	6.12 (1.23)	6.80 (1.36)
$\sigma_3 = 2.0$	6.84 (1.43)	6.88 (1.31)	6.12 (1.23)	6.55 (1.40)
$n_1/n = 0.04$	7.02 (1.35)	7.08 (1.23)	6.12 (1.23)	6.77 (1.40)
$n_1/n = 0.20$	8.65 (0.96)	8.66 (0.96)	8.25 (1.00)	8.52 (1.00)
$n_1/n = 0.36$	9.22 (0.82)	9.23 (0.81)	9.04 (0.83)	9.13 (0.83)
$n_1/n = 0.50$	9.51 (0.72)	9.51 (0.72)	9.39 (0.71)	9.43 (0.72)

Table 3.4: Estimated value function of B_2 (mean (sd)) based on SS-learning with proposed tuning, SS-learning with oracle tuning, OWL on $\mathcal{S}_1$, and OWL on $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$.

	SS-learning (tune)	SS-learning (oracle)	OWL on $\mathcal{S}_1$	OWL on $\mathcal{S}$
$\beta_3 = 0.1$	1.34 (0.24)	1.35 (0.23)	1.31 (0.23)	1.33 (0.23)
$\beta_3 = 0.5$	1.36 (0.23)	1.36 (0.23)	1.31 (0.23)	1.35 (0.23)
$\beta_3 = 1.0$	1.36 (0.23)	1.37 (0.23)	1.31 (0.23)	1.36 (0.24)
$\beta_3 = 3.0$	1.36 (0.25)	1.37 (0.23)	1.31 (0.23)	1.36 (0.23)
$\sigma_3 = 0.1$	1.36 (0.26)	1.37 (0.24)	1.31 (0.23)	1.36 (0.23)
$\sigma_3 = 0.5$	1.36 (0.24)	1.36 (0.23)	1.31 (0.23)	1.36 (0.23)
$\sigma_3 = 1.0$	1.36 (0.24)	1.36 (0.23)	1.31 (0.23)	1.35 (0.23)
$\sigma_3 = 2.0$	1.34 (0.26)	1.36 (0.23)	1.31 (0.23)	1.34 (0.23)
$n_1/n = 0.04$	1.35 (0.25)	1.36 (0.24)	1.31 (0.23)	1.35 (0.23)
$n_1/n = 0.20$	1.32 (0.22)	1.32 (0.23)	1.31 (0.23)	1.32 (0.23)
$n_1/n = 0.36$	1.36 (0.23)	1.36 (0.22)	1.35 (0.23)	1.35 (0.23)
$n_1/n = 0.50$	1.36 (0.24)	1.37 (0.24)	1.36 (0.24)	1.36 (0.24)

Table 3.5: Estimated value function of B_3 (mean (sd)) based on SS-learning with proposed tuning, SS-learning with oracle tuning, OWL on $\mathcal{S}_1$, and OWL on $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$.

	SS-learning (tune)	SS-learning (oracle)	OWL on $\mathcal{S}_1$	OWL on $\mathcal{S}$
$\beta_3 = 0.1$	3.06 (0.38)	3.06 (0.38)	3.01 (0.39)	3.05 (0.38)
$\beta_3 = 0.5$	3.53 (0.42)	3.53 (0.42)	3.31 (0.43)	3.46 (0.43)
$\beta_3 = 1.0$	4.56 (0.45)	4.56 (0.44)	4.10 (0.52)	4.50 (0.44)
$\beta_3 = 3.0$	9.90 (1.05)	9.96 (0.78)	8.49 (1.01)	9.85 (0.80)
$\sigma_3 = 0.1$	1.60 (0.17)	1.61 (0.13)	1.37 (0.17)	1.53 (0.17)
$\sigma_3 = 0.5$	2.30 (0.24)	2.30 (0.24)	2.07 (0.27)	2.23 (0.27)
$\sigma_3 = 1.0$	3.54 (0.38)	3.54 (0.38)	3.33 (0.41)	3.48 (0.39)
$\sigma_3 = 2.0$	6.28 (0.86)	6.30 (0.72)	6.11 (0.73)	6.23 (0.71)
$n_1/n = 0.04$	3.57 (0.45)	3.59 (0.38)	3.37 (0.40)	3.52 (0.39)
$n_1/n = 0.20$	3.49 (0.43)	3.49 (0.43)	3.41 (0.44)	3.46 (0.43)
$n_1/n = 0.36$	3.41 (0.44)	3.41 (0.44)	3.38 (0.44)	3.40 (0.44)
$n_1/n = 0.50$	3.25 (0.48)	3.26 (0.48)	3.23 (0.48)	3.24 (0.47)

objective of our analysis is to derive an ITR for pregnant women in taking calcium supplementation to maximize the reduction of fetal blood lead concentration at birth. The primary outcome B_1 is the cord blood lead concentration, the second outcome B_2 is the blood lead concentration of infant at 3-month age, and the third outcome is the maternal blood lead concentration at the third trimester. SS-learning is based on an ordering benefit value, with the larger magnitude of benefits, the more clinical desirable outcome. To fit in this framework, we reversed the direction of lead concentration by a transformation, $\max_i(B_i) - B_i, i = 1, \ldots, n$. After deleting observations with missing data, 59 participants have cord blood lead measurements (37 taking calcium supplement; 22 taking placebo), 190 participants have infant 3-month blood lead concentration measurements (95 taking calcium supplement; 95 taking placebo), and 116 participants have maternal blood lead concentration measurements at the third trimester (51 taking calcium supplement; 65 taking placebo). We performed an intention-to-treat analysis under the missing completely at random (MCAR) assumption, since the attrition of measurements was due mainly to the logistic issues of biosample collection not related to lead exposure. Since the number of subjects

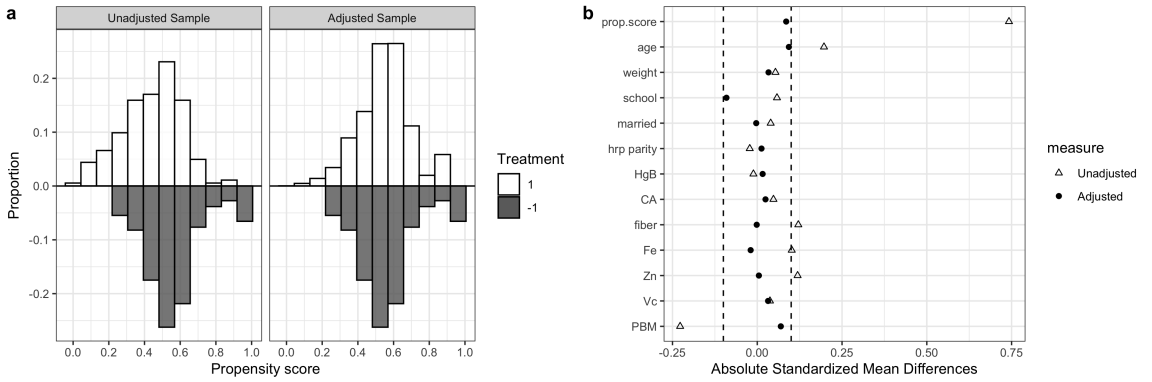

Figure 3.4: Evaluation of predictor balance. (a) Distributional balance of propensity score before and after weighting. (b) predictor balance before and after weighting illustrated by the absolute standardized mean differences. The thresholds -0.1 and 0.1 are shown as the vertical dashed lines.

with lead concentration measurements from the two arms are not fully balanced (as is shown in Table 3.1), the propensity score $P(A \mid \boldsymbol{X})$ was estimated by logistic regression based on all the major effects of the predictors.

After weighting the observations by the estimated propensity score, we evaluated the balance of the propensity score by plotting the group-specific distributions and the balance of the predictors using the absolute standardized mean differences (see Figure 3.4). It is shown that predictor balance has been met within a threshold of 0.1, and that balance has improved after weighting for all the predictors, as well as for the propensity score.

We then came to the major objective to derive an ITR for calcium supplementation to pregnant women based on the predictors. We performed: (i) a standard OWL analysis on the dataset $\mathcal{S}_1$ with B_1 and got the estimated decision rule $\hat{f}_1$; (ii) a standard OWL analysis on the augmented training data by including $\mathcal{S}_2$ with B_2 and $\mathcal{S}_3$ with B_3, and naively ignored the differences of clinical relevance (i.e. $\lambda_2 = \lambda_3 = 1$) and got the estimated decision rule $\hat{f}_2$; (iii) the proposed SS-learning on the whole dataset in which we took into account differences between B_1, B_2 and B_3 in terms of their clinical relevance to the underlying health benefit and got the estimated decision

rule $\hat{f}_3$. Linear kernel was used for solving the optimization function (see Algorithm III.1) of SS-learning and linear regressions were applied for parameter tuning (see Section 3.3.3). The resulting estimated decision rules are given as

$$\begin{aligned}
\hat{f}_1 =& 0.02 + 0.47\text{age} - 0.03\text{weight} + 0.03\text{school} + 0.72\text{married} + 0.19\text{parity} \\
& - 0.66\text{HgB} - 0.33\text{Ca} + 0.03\text{fiber} - 0.10\text{Fe} + 0.48\text{Zn} + 0.09\text{VC} - 0.01\text{PBM} \\
\hat{f}_2 =& -0.20 + 0.18\text{age} - 0.12\text{weight} + 0.22\text{school} - 0.05\text{married} + 0.03\text{parity} \\
& - 0.05\text{HgB} - 0.33\text{Ca} + 0.61\text{fiber} + 0.04\text{Fe} + 0.05\text{Zn} - 0.04\text{VC} - 0.31\text{PBM} \\
\hat{f}_3 =& -0.33 + 0.57\text{age} - 0.17\text{weight} - 0.08\text{school} + 0.60\text{married} + 0.11\text{parity} \\
& + 0.26\text{HgB} - 0.81\text{Ca} + 0.67\text{fiber} + 0.17\text{Fe} + 0.21\text{Zn} + 0.03\text{VC} - 0.57\text{PBM}
\end{aligned}$$

Using the proposed parameter tuning procedure, we obtained $\lambda_2 = 0.8$ and $\lambda_3 = 0.3$ in the estimation of $\hat{f}_3$, indicating that infant 3-month blood lead concentration has slightly lower clinical relevance, while maternal blood lead concentration at the third trimester has much lower clinical relevance, than cord blood lead concentration when used to derive the ITR. Applying the estimated decision rules on each subject, we would allocate a pregnant woman to take the calcium supplement if $\hat{f}_q > 0, q = 1, 2, 3$; otherwise, she does not need to do so because of no benefit. Among the total 365 individuals, $\hat{f}_1, \hat{f}_2$ and $\hat{f}_3$ assigns 277, 130, and 210 subjects to take the calcium supplement, respectively. Comparing the treatment assignment results between complete randomization A and the estimated decision rule $\hat{f}_3$ by SS-learning, we found a significantly different reallocation (McNemar test p-value 0.036) (see Table 3.6), indicating that the difference in treatment allocation does not arise by accident.

The SSE values given by $\hat{f}_1, \hat{f}_2$ and $\hat{f}_3$ are 1365.6, 1355.2 and 1342.2, indicating that the linear regressions derived from $\hat{f}_3$ give the best prediction of the benefit

Table 3.6: Treatment assignment comparison between complete randomization and $\hat{f}_3$ estimated by SS-learning.

		$\hat{f}_4$		
		placebo	calcium	total
randomization	placebo	86	96	182
	calcium	69	114	183
	total	155	210	365

Table 3.7: Estimated value functions of B_1, B_2 and B_3 based on A, the estimated ITR $\hat{f}_1, \hat{f}_2, \hat{f}_3$ for the analysis of the calcium supplementation trial.

	$\hat{V}(B_1)$	$\hat{V}(B_2)$	$\hat{V}(B_3)$
A	8.80	8.86	10.92
$\hat{f}_1$	8.79	8.98	11.37
$\hat{f}_2$	8.70	8.85	10.97
$\hat{f}_3$	8.97	8.98	11.45

values. In order to illustrate the performance of $\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$ in reducing the maternal lead exposure to infant, we also calculated the estimated value functions of B_1, B_2 and B_3 as $\mathbb{E}_n^*\{I(A = D(\boldsymbol{X}))B/P(A \mid \boldsymbol{X})\}/\mathbb{E}_n^*\{I(A = D(\boldsymbol{X}))/P(A \mid \boldsymbol{X})\}$ (*Murphy et al.*, 2001), where $\mathbb{E}_n^*$ denotes the empirical average. The estimated value functions based on complete treatment randomization and the three estimated decision rules are summarized in Table 3.7. It shows that $\hat{f}_3$ gives the highest estimated value functions for B_1, B_2 and B_3, suggesting the largest reduction in fetal blood lead instead of $\hat{f}_1$ or $\hat{f}_2$ when they were implemented for the whole study population. This indicates that the inclusion of B_2 and B_3 by taking the heterogeneous clinical relevance into account has improved the estimated ITR in terms of the estimated benefit values.

Checking the demographic and baseline predictors of the individuals (see Table 3.8), we found that subjects assigned to take calcium supplementation by $\hat{f}_3$ have a significantly older age in agreement with the positive coefficient of predictor *age* in $\hat{f}_3$. That is, older pregnant women are more likely to benefit from calcium supplement. In addition, subjects assigned to taking calcium supplement by $\hat{f}_3$ have a significantly higher proportion of marriage, a significantly larger hrp parity, a signif-

Table 3.8: Summary statistics of the predictors based on treatment allocation according to $\hat{f}_3$. Mean (sd) and percentage values are shown for the numeric and categorical predictors respectively. p-values are obtained from t-test and chi-aquare test.

Predictor	Calcium	Placebo	p-value
age	28.5 (5.2)	23.7 (4.9)	$< 2.20 \times 10^{-16}$
weight	62.4 (10.7)	61.3 (10.2)	2.89×10^{-1}
school	11.0 (3.1)	10.4 (2.8)	5.56×10^{-2}
married	0.771	0.581	1.55×10^{-4}
hrp	2.3 (1.0)	1.9 (1.0)	1.56×10^{-3}
HgB	13.4 (0.8)	13.0 (1.2)	3.36×10^{-4}
Ca	10.4 (4.8)	11.6 (5.6)	2.55×10^{-2}
fiber	26.3 (9.6)	19.8 (8.0)	8.89×10^{-12}
Fe	14.4 (5.9)	11.7 (5.2)	7.53×10^{-6}
Zn	9.8 (3.4)	9.0 (3.8)	3.13×10^{-2}
VC	192.1 (90.5)	150.9 (73.9)	2.61×10^{-6}
PBM	4.2 (2.7)	6.1 (3.6)	1.08×10^{-7}

icantly higher HgB concentration, a significantly lower dietary intake of calcium, a significantly higher dietary intake of fiber, Fe, Zn and vitamin C, and a significantly lower baseline blood lead concentration. These are also aligned with their coefficients in $\hat{f}_3$, indicating that pregnant women of being married, being pregnant for more times, having higher HgB concentration, having lower calcium intake from food, having higher dietary intake of fiber, Fe, Zn, and vitamin C and having lower level of baseline blood lead tend to benefit from calcium supplement. Interestingly, subjects assigned to the calcium supplement by $\hat{f}_3$ have a lower calcium intake from food (1037.3 (482.5) versus 1163.5 (563.7)), meaning that participants who have enough dietary calcium do not need to take calcium supplement. This finding is clinically appealing. However, no significant differences were identified for weight and years in school between the two allocated groups.

3.6 Concluding Remarks

In this chapter, we derived an ITR for pregnant women in the take of daily calcium supplementation to maximize the reduction of maternal lead exposure to infants during pregnancy using SS-learning based on the ELEMENT calcium trial. The major difficulty in analyzing the data lies in the fact that we have to integrate multiple outcomes of interest with heterogeneous clinical relevance and sample sizes in the derivation of ITR, in order to achieve high reproducibility and robustness of the derived ITR. SS-learning solves this problem based on a weighted self-learning paradigm with characterization of the relative contributions of respectively low- and high-quality data sources. SS-learning is particularly appealing in ITR derivation using real clinical trial data when measuring high quality data of high clinical relevance is often expensive, resulting in a small sample size, while more data sources of much larger sample sizes but likely of low clinical relevance are available. Simulation studies in Section 3.4 show that our method converges fast and outperforms the existing decision rule estimation method OWL.

To our knowledge, it is the first time such an ITR in assigning calcium supplementation to pregnant women is established. The results reveal that the auxiliary data sources of infant 3-month blood lead concentration and maternal blood lead concentration at the third trimester have a lower clinical relevance than the gold cord blood lead concentration in the derivation of ITR for calcium supplementation allocation. But the integration of these secondary outcomes of interest do improve the derived ITR with respect to the expected lead reduction compared to the ITR estimated by the primary outcome alone or by naively treating the different outcomes as having equal clinical relevance. The resulting rule can increase lead exposure reduction in uterus. Clinically, it is of great importance to improve a decision rule given for further reduction of blood lead concentration, which has been accomplished by our analysis based on the proposed SS-learning method. In addition to the different clinical rele-

vance we have characterized for the three outcomes used for ITR estimation, we also identified the variables that play significant roles in allocating calcium supplementation to pregnant women. Interestingly, dietary intake of fiber, Fe, Zn and vitamin C will influence the assignment of calcium supplement, indicating a complex network of interactions involving different nutrient contents. This suggests us to consider more comprehensively when making recommendations for individuals to take nutritional supplements.

In this article, we employed linear kernel to derive the decision rule $\hat{f}_3$ using SS-learning to easily illustrate the derived ITR. Nonlinear-kernels can also be applied to derive more complicated decision rules. Another simulation experiment included in the online Supplementary Materials shows the performance of radial basis function (RBF) kernel in the establishment of ITR. Different regression models are also compared to get the tuning parameters. In the application of SS-learning, selection of kernel functions and parameter tuning models are of great importance to guarantee good performance of the derived ITR. Since the linear ITR derived by SS-learning already gives higher expected benefits compared to the decision rules estimated using traditional OWL for this ELEMENT calcium supplementation trial, we did not try nonlinear kernels, which can be a further direction for future analysis. Additional future research can be devoted to extending this methodology to clinical trial with multiple treatment arms, to situations where there are categorical outcomes or time-to-event outcomes, and to dynamic treatment regimes. Moreover, we may alleviate the computing burden of grid search in the tuning when the number of auxiliary datasets is large by grouping auxiliary datasets with similar quality and relevance, so to use the same λ value for each group, or by performing target search of λ value within a certain smaller range.

CHAPTER IV

Longitudinal Self-Learning of Individualized Treatment Rules with Missing Data

4.1 Introduction

This chapter presents an exemplary applied statistics contribution to the emerging field of precision nutrition. As is mentioned in Chapter III, Barker's DOHaD hypothesis postulates a key conceptual paradigm for effects of perinatal environment on the future health of offspring. We are still interested in controlling the maternal lead exposure to infants, since a vast literature has unveiled that excessive exposure to lead is detrimental on children's neurobehavioral and cognitive development (*Chen et al.*, 2005; *Hornung et al.*, 2009; *Wasserman et al.*, 2003; *Téllez-Rojo et al.*, 2006b; *Hu et al.*, 2006; *Braun et al.*, 2012). As is introduced in Chapter III, reducing the amount of lead released into maternal blood circulation during pregnancy is a significant preventive measure to minimize lead exposure to children, as prenatal lead exposure by fetus comes solely from mother. Clinically, blood lead control may be achieved by calcium.

In this study of precision nutrition, we are interested in establishing an individualized treatment rule (ITR) that can be tailored to pregnant women in taking daily calcium supplement with the aim to reduce lead exposure to children of 3-years old.

Not all pregnant women will have a long-term benefit from taking calcium supplementation, which may due to two major reasons: one is that some pregnant women already have enough calcium intake from their dietary food, and the other is that too high blood calcium may increase the risk of miscarriage (*Norman et al.*, 2009). The derivation of such an ITR usually targets at the maximization of the expected outcome. In this chapter, we still focus on the calcium supplementation trial from the Early Life Exposure in Mexico to ENvironmental Toxicants (ELEMENT) Project. But different from that in Chapter III, in particular, the endpoint outcome of interest in this analysis is blood lead concentration for children (PBC; "PB" represents lead and "C" represents children) at 36-month of age. This longitudinal trial collects repeated measurements of PBC values at month 3, 6, 12, 18, 24, 30 and 36. We treat the outcome measured at month 36 (denoted by PBC36) as the primary endpoint for ITR derivation. Due to intermittent missing data or dropouts, the observations of the PBC36 are incomplete (36.7% missingness), leading to the sample size attrition if only the endpoint outcome is used. Consequently, both reliability and reproducibility of the resulting ITR would be compromised. Therefore, one natural solution is to utilize longitudinal data of blood concentration prior to PBC36 to construct ITR, where a certain temporal weighting scheme is specified to reflect their clinical relevance to the endpoint outcome.

Statistical methodologies for estimating decision functions in the field of precision medicine are abundant, such as outcome weighted learning (OWL) (*Zhao et al.*, 2012). Unfortunately, all the current learning methods are applicable only for a one-dimensional cross-sectional outcome related to health benefit. It lacks suitable methods in the literature that utilize temporally correlated outcomes and handle missing data in both estimation and optimization of ITRs. To analyze the longitudinal calcium supplementation trial, we extend the SS-learning method proposed in Chapter III to longitudinal self-learning (LS-learning), which is flexible to not only

incorporate longitudinal outcomes but also to handle incomplete data in ITR derivation. LS-learning works in a weighted self-learning paradigm via an effective training data augmentation scheme, which is built upon the assumption that the relevance of longitudinal outcomes increases over time towards the primary endpoint. Moreover, estimated subgroup labels learned from longitudinal outcomes measured prior to the primary endpoint are iteratively calibrated with those labels determined by the endpoint. Same as that for SS-learning, the relative contributions from longitudinal outcomes are tuned by minimizing the sum of squared errors (SSE) of the predicted treatment benefits. To handle missing data, we follow the idea of pattern mixture model (*Little*, 2008) that stratifies subjects into subgroups according to their missing data patterns. To deal with potential computational burden in tuning parameter selection as is discussed in Section 3.6, we adopt a scalable tuning procedure incorporating the nature of longitudinal data collection. Meanwhile, using extensive simulation experiments, we illustrate and confirm that the LS-learning algorithm can numerically achieve both high classification rate for treatment subgroups and high computational efficiency.

We apply LS-learning to derive an ITR in the longitudinal calcium supplementation trial, where we find that LS-learning would produce the largest benefit in reducing children's exposure to blood lead concentration at 36 months of age if the resulting ITR were implemented for the whole study population of pregnant women, in comparison to other ITRs derived from standard OWL. This improvement is of clinical importance and is achieved by the means of longitudinal data augmentation in our analysis, which showcases an exemplary approach to deliver better solutions in precision nutrition. Several predictors that play an important role in forming the ITR are also identified. Interestingly, dietary intake of fiber, Fe, Zn and vitamin C will influence the assignment of calcium supplement, indicating a complex group of interactions involving different nutrient contents.

The remainder of the chapter is organized as follows. Section 4.2 introduces the longitudinal calcium supplementation trial. Section 4.3 presents the LS-learning approach, including its systematic parameter tuning method. Section 4.4 concerns the evaluation of LS-learning through simulation experiments. Data analysis for the longitudinal calcium trial is detailed in Section 4.5. Section 4.6 contains some concluding remarks.

4.2 Application: Longitudinal Calcium Supplementation Trial

In the ELEMENT Project, one of the major objectives is to assess how environmental exposures, such as heavy metals or endocrine disrupting chemicals (EDCs), affect the health outcomes of pregnant women and their children (*Téllez-Rojo et al.*, 2004; *Zhang et al.*, 2012). This chapter still focuses on the calcium supplementation trial conducted in the third cohort of the ELEMENT study.

As is introduced in Section 3.2, this clinical trial contains a total of 670 women who were recruited at their first trimester of pregnancy during 2001 to 2003 in Mexico City, with 334 mothers randomized to receive a daily dose of 1200 mg calcium supplement and 336 randomized to receive placebo (*Zhang et al.*, 2012). These mothers are followed during both pre- and post-natal pregnancy. Demographic information and biological samples from both mothers and children (i.e. mother-child pairs) are collected for clinical measurements. The primary outcome of interest is child's blood lead concentration (PBC) at 36-month of age, denoted by PBC36, which is an endpoint of clinical importance due to its known influence on children neurobehavioral development (*Wasserman et al.*, 1998; *Mendelsohn et al.*, 1998). In addition to this primary PBC36 outcome, six repeated measurements of PBC at month 3, 6, 12, 18, 24, and 30 are collected during the followup visits as part of the longitudinal study, which for convenience are denoted by PBC3, PBC6 and so on. Figure 4.2 displays the longitudinal trajectories of PBC values in the calcium supplementation and placebo

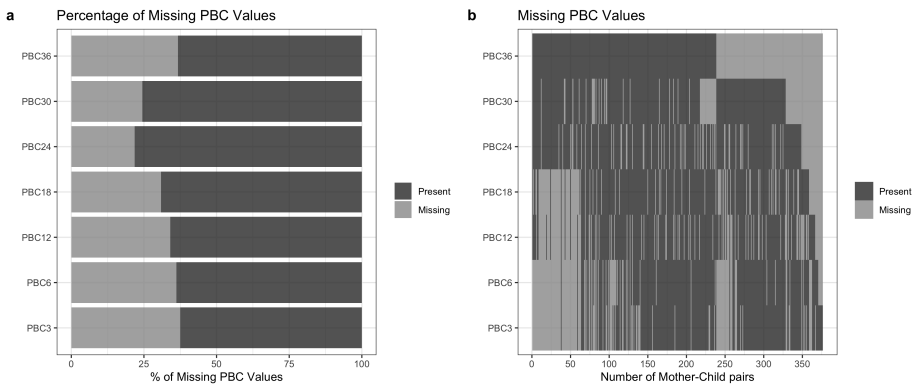

Figure 4.1: Missing of PBC values at different visit times. (a) Percentage of missing and observed PBC values at month 3, 6, 12, 18, 24, 30, and 36. (b) Stratification of mother-child pairs into different missing patterns based on individual endpoints.

group. Instead of deleting mother-child pairs with missing measurement of PBC36 (37.6% missingness), we intend to borrow longitudinal PBC values measured prior to month 36 into ITR derivation, resulting in an improved ITR by recovering data attrition. We propose to stratify mother-child pairs in terms of their missing data patterns, leading to subgroups with specific last available observation, which is taken as a surrogate or an approximate endpoint of lead exposure measured at an earlier time point than month 36. Figure 4.1 shows the percentage of missingness for different PBC values and the stratification of subgroups with different individual endpoints. For example, if a child has PBC observed at month 3, 12, 18, and 30, the last available PBC observation of this individual, namely PBC30, is taken as a surrogate endpoint which will be used in ITR derivation.

We choose the same group pf baseline predictors as those in Chapter III to be used for ITR derivation. Summary statistics of these predictors in the calcium supplementation and placebo group, as well as the differences of their distributions characterized by t-test or chi-square test are listed in Table 4.1. The final training dataset excludes 293 mother-child pairs that have missing data in the baseline predictors. The data we received from the ELEMENT group is the one with all baseline predictors fully

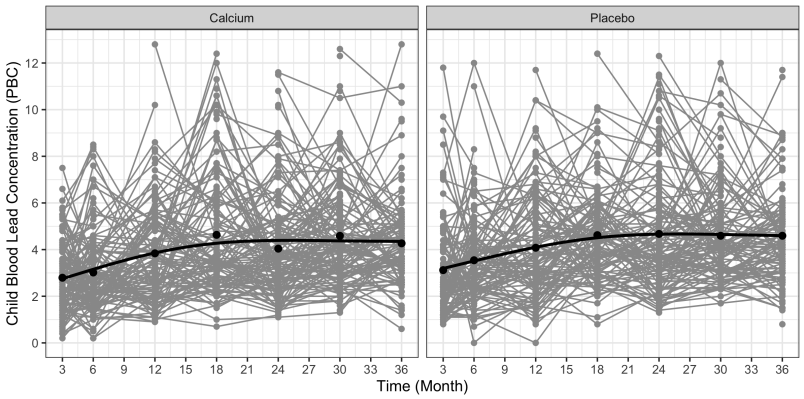

Figure 4.2: Trajectories of PBC values in the calcium supplementation and placebo group. The black summary line is fitted by the generalized additive model (GAM).

Table 4.1: Summary statistics of the predictors included in ITR derivation. Mean (sd) and percentage are shown, where p-values are obtained from t-test and chi-square test for numeric and categorical variables, respectively.

Predictor	Calcium	Placebo	p-value
age	26.9 (5.7)	25.9 (5.4)	0.071
weight	62.0 (11.3)	61.6 (9.6)	0.717
school	10.8 (2.9)	10.7 (2.9)	0.600
married	0.695	0.677	0.802
hrp parity	2.1 (1.0)	2.1 (1.1)	0.867
HgB	13.2 (1.0)	13.2 (1.0)	0.800
Ca	11.2 (4.9)	10.8 (5.5)	0.493
fiber	24.4 (10.0)	22.8 (9.0)	0.125
Fe	13.7 (5.9)	12.9 (5.5)	0.223
Zn	9.8 (3.5)	9.3 (3.6)	0.173
VC	178.1 (92.1)	172.7 (80.4)	0.548
PBM	4.7 (2.7)	5.3 (3.7)	0.070

observed. In addition, we further exclude one mother-child pair with all longitudinal PBC values missing at the 7 visits.

4.3 Formulation of LS-learning

This section presents the details of longitudinal self-learning (LS-learning) that will be used to derive an ITR using longitudinal data with missing outcomes. This LS-learning helps us yield a novel ITR from the calcium supplementation trial to reduce continuous lead exposure for children at 36-month of age.

4.3.1 Notation

Denote the longitudinal data $\mathcal{S} = \{(\boldsymbol{X}_i, A_i, t_{ij}, B_{ij}, R_{ij}), i = 1, \ldots, n, j = 0, \ldots, m_i\}$ collected from a two-armed randomized clinical trial (i.e. the longitudinal calcium supplementation trial). Let i and j denote the index of subject (i.e. mother-child pair i) and followup visit j, respectively. These two indices i or j may be suppressed when there is no confusion. In the rest of this chapter, "subject" represents "mother" when it comes to the measurements of baseline predictors, or represents "child" when it refers to the measurements of outcomes (i.e. the longitudinal PBC values), or represents "mother-child" pair when data from both are in the use. There is a total of n mother-child pairs and m_i repeated measurements for subject i, and $j = 0$ corresponds to the baseline visit prior to treatment randomization. For ease of exposition, we consider a longitudinal trial with outcomes measured at common visit times $j = 0, \ldots, m$, where $j = m$ stands for the time for the measurement of primary endpoint (i.e. PBC36, which is the PBC value at the 36-month visit when $j = 7$). Consequently, visit times $t_{ij} \equiv t_j, j = 1, \ldots, m$. Assigned treatment at the beginning of the trial is denoted as $A_i \in \mathcal{A} = \{-1, 1\}$, where $A_i = 1$ represents the new treatment (i.e. calcium supplementation) and $A_i = -1$ represents placebo. $\boldsymbol{X}_i \in \mathbb{R}^d$ is a d-dimensional vector of predictors measured at baseline. The outcome of interest,

denoted by B_{ij}, is repeatedly measured at time t_j for subject i. It is assumed that the larger value of B the higher benefit to health. Specifically, according to the clinical study, it is the outcome B_m that serves as the primary outcome to be used for ITR derivation. All the other intermittent measurements, $B_1, \ldots, B_{m-1}$, are treated as surrogate outcomes providing relevant and auxiliary information of benefit trajectory reaching the endpoint B_m. Due to dropouts or other reasons, some of the individuals in the trial are not measured at the final visit t_m, resulting in missing data of the primary endpoint B_m. For those subjects that do not complete the trial, instead of deleting them from the analysis, certain surrogate endpoints will be chosen so that they can remain in our ITR derivation. Let $R_{ij} \in \{0, 1\}$ denote the missingness of outcome B_{ij}, with $R_{ij} = 1$ representing the missing value and $R_{ij} = 0$ representing the observed value.

4.3.2 Longitudinal Self-learning

As is discussed in Chapter III, SS-learning is proposed to overcome the limitation that OWL only works for a single cross-sectional outcome. In this chapter, we further propose LS-learning, which is an extension to SS-learning that addresses the methodological need in ITR deviation with longitudinal outcomes. We proceed this extension in the setting of multiple training datasets, obtained specifically from stratification on missing-data patterns in a clinical trial with dropouts. Denote $\mathcal{S}_m = \{(\boldsymbol{X}_i, A_i, t_m, B_{im}), i = 1, \ldots, n_m\}$ as the subset of completers, namely the subjects who have measurements B_m. All the other subjects $i \in \bar{\mathcal{S}}_m$ have their individual surrogate endpoints, depending on specific available last observations over dropout time points. This results in $m - 1$ subsets $\mathcal{S}_j = \{(\boldsymbol{X}_i, A_i, t_j, B_{ij}), i = 1, \ldots, n_j\}, j \in \mathcal{J} = \{1, \ldots, m - 1\}$, each corresponding to one stratum. Obviously, some of subsets may be empty; for example, all of them are empty sets if all subjects complete the trial. To estimate ITR, we plan to apply LS-learning that integrates the subset of

completers $B_{im}, i \in \mathcal{S}_m$ with all the other subsets $B_{ij}, i \in \mathcal{S}_j, j \in \mathcal{J}$ in a systematic way. The proposed LS-learning algorithm in a generic fashion is given as follows.

Algorithm IV.1. Suppose we have a primary dataset $\mathcal{S}_m = \{(\boldsymbol{X}_i, A_i, t_m, B_{im}), i = 1, \ldots, n_m\}$ with the outcome taken at t_m and multiple auxiliary datasets $\mathcal{S}_j = \{(\boldsymbol{X}_i, A_i, t_j, B_{ij}), i = 1, \ldots, n_j\}, j \in \mathcal{J} = \{1, \ldots, m-1\}$ with the outcome taken prior to t_m. Assume that B_m is of higher quality than B_j in ITR derivation. Let $p_i = P(A_i \mid \boldsymbol{X}_i)$ be the propensity score. We iterate the following steps.

S1 Estimate ITR by OWL using $\mathcal{S}_m$ and get the initial estimates $(\boldsymbol{\omega}^{(0)}, b^{(0)}, \boldsymbol{\xi}^{(0)})$. Then predict labels for subjects in $\bigcup_{j=1}^{m-1} \mathcal{S}_j$. Denote the predicted labels as $y_{ij}^{(0)}, i \in \mathcal{S}_j, j \in \mathcal{J}$.

S2 The k-th ($k \geq 1$) iteration runs through the following steps S2.1-S2.3.

S2.1 Define an augmented training dataset $\mathcal{S}^{(k)} = \mathcal{S}_m \cup \tilde{\mathcal{S}}_1^{(k)} \cup \cdots \cup \tilde{\mathcal{S}}_{m-1}^{(k)}$, where individuals in $\tilde{\mathcal{S}}_j^{(k)}$ have their predicted labels from the previous iteration $k-1$, resulting in $\tilde{\mathcal{S}}_j^{(k)} = \{(\boldsymbol{X}_i, A_i, t_j, B_{ij}, y_{ij}^{(k-1)}), i = 1, \ldots, n_j\} = \mathcal{S}_j \cup \{y_{ij}^{(k-1)}, i = 1, \ldots, n_j\}, j \in \mathcal{J}$.

S2.2 Solve the following optimization problem using $\mathcal{S}^{(k)}$:

$$
\begin{aligned}
\min_{\boldsymbol{\omega}, b, \boldsymbol{\xi}} \quad & \frac{1}{2}||\boldsymbol{\omega}||^2 + C\left[\sum_{i \in \mathcal{S}_m} \frac{B_{im}}{p_i}\xi_i + \sum_{j=1}^{m-1}\left\{\lambda_j \sum_{i \in \tilde{\mathcal{S}}_j^{(k)}} \frac{B_{ij}}{p_i}\widetilde{\xi}_{ij} + (1-\lambda_j)\sum_{i \in \tilde{\mathcal{S}}_j^{(k)}} \tau_{B_j}\check{\xi}_{ij}\right\}\right] \\
\text{s.t.} \quad & A_i(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0,\ i \in \mathcal{S}_m, \\
& A_i(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \widetilde{\xi}_{ij} \text{ and } \widetilde{\xi}_{ij} \geq 0,\ i \in \tilde{\mathcal{S}}_j^{(k)},\ j \in \mathcal{J}, \\
& y_{ij}^{(k-1)}(\boldsymbol{\omega}^\intercal \boldsymbol{X}_i + b) \geq 1 - \check{\xi}_{ij} \text{ and } \check{\xi}_{ij} \geq 0,\ i \in \tilde{\mathcal{S}}_j^{(k)},\ j \in \mathcal{J}.
\end{aligned} \tag{4.1}
$$

where $\lambda_j \in [0, 1]$ is an additional tuning parameter that characterizes the relative contribution of the predicted and observed information from $\tilde{\mathcal{S}}_j^{(k)}$. τ_{B_j} is a summary statistic (e.g. sample mean) of $\frac{B_{ij}}{p_i}$ that is added as the

weight for $\check{\xi}_{ij}$ to balance the magnitude between $\sum_{i\in\tilde{\mathcal{S}}_j^{(k)}} \widetilde{\xi}_{ij}$ and $\sum_{i\in\tilde{\mathcal{S}}_j^{(k)}} \check{\xi}_{ij}$, so to reduce the bias over the tuning of λ_j. Denote the estimates of the parameters as $(\boldsymbol{\omega}^{(k)}, b^{(k)}, \boldsymbol{\xi}^{(k)})$ at iteration k. Also, the predicted labels for $\tilde{\mathcal{S}}_j^{(k)}$ are updated as $y_{ij}^{(k)}, i \in \tilde{\mathcal{S}}_j^{(k)}, j \in \mathcal{J}$.

S2.3 Calculate the objective function at the estimates $(\boldsymbol{\omega}^{(k)}, b^{(k)}, \boldsymbol{\xi}^{(k)})$ as

$$h(\boldsymbol{\omega}^{(k)}, \boldsymbol{\xi}^{(k)}) = \frac{1}{2}||\boldsymbol{\omega}^{(k)}||^2 + C\left[\sum_{i\in\mathcal{S}_m} \frac{B_{im}}{p_i}\xi_i^{(k)} + \sum_{j=1}^{m-1}\left\{\lambda_j \sum_{i\in\tilde{\mathcal{S}}_j^{(k)}} \frac{B_{ij}}{p_i}\widetilde{\xi}_{ij}^{(k)} + (1-\lambda_j)\sum_{i\in\tilde{\mathcal{S}}_j^{(k)}} \tau_{B_j}\check{\xi}_{ij}^{(k)}\right\}\right]$$

S3 The algorithm stops if $\frac{|h(\boldsymbol{\omega}^{(k)},\boldsymbol{\xi}^{(k)})-h(\boldsymbol{\omega}^{(k-1)},\boldsymbol{\xi}^{(k-1)})|}{h(\boldsymbol{\omega}^{(k-1)},\boldsymbol{\xi}^{(k-1)})} < \epsilon$ for a pre-determined precision constant ϵ, say 10^{-4}. The convergence values of $(\boldsymbol{\omega}^{(k)}, b^{(k)})$ are denoted by $(\hat{\boldsymbol{\omega}}, \hat{b})$, and the predicted labels at convergence for the whole training data $\mathcal{S}$ are denoted as $\hat{y}_i, i = 1, \ldots, n$.

The proposed Algorithm IV.1 is similar to Algorithm III.1 of SS-learning in Chapter III. The major difference lies in the primary benefit outcome used for ITR derivation, which is the primary endpoint outcome B_m for LS-learning and the most high-quality benefit outcome B_1 in SS-learning. Meanwhile, LS-learning stratifies subjects into subgroups with different individual endpoint outcomes according to their missing data patterns, which is a procedure that is not employed in SS-learning.

4.3.3 Scaled Tuning Scheme

To perform LS-learning, we need to determine the tuning parameters C and $\boldsymbol{\lambda} = \{\lambda_j, j = 1, \ldots, m-1\}$. As is proposed in Chapter III for SS-learning, parameter selection may be done through a grid search of C and $\boldsymbol{\lambda}$ on the entire dataset $\mathcal{S}$ according to the minimization of the sum of squared errors (SSE) for the pre-

dicted benefit values. SSE is defined as $\text{SSE} = \sum_{j=1}^{m} \left\{ \sum_{i=1}^{n_j} (B_{ij} - \hat{B}_{ij})^2 \right\}$, where the predicted benefit value $\hat{B}_{ij}$ is obtained from the stratum-specific regression model $B_j = \phi_j(\boldsymbol{X}, \boldsymbol{s}(t_j), A\hat{f}(\boldsymbol{X})) + \epsilon_j$, with $\phi_j(a, b, c) = g_j(a) + h_j(b) + \beta_j c$ for some function $g_j(\cdot)$ and $h_j(\cdot)$. $\hat{f}(\cdot)$ is the estimated decision function via LS-learning and ϵ_j is the error term. Multiple linear regression model or more flexible models such as the generalized additive model (GAM) may be invoked to build the prediction rule $\phi_j(\cdot)$. In the simulation experiment described in Section 4.4, as well as the calcium trial analysis discussed in Section 4.5, we choose GAM for ϕ_j to predict the longitudinal benefit outcomes.

Each auxiliary dataset $\mathcal{S}_j$ will introduce a λ_j to tune, and the strategy of greedy search proposed for SS-learning may make parameter tuning computationally expensive and time consuming. To alleviate this computational burden, we propose a scaled tuning method that specifies λ_j as a function of the standardized time $\tilde{t}_j = t_j/t_m$, i.e. $\lambda_j = \psi(\tilde{t}_j \mid \boldsymbol{\theta})$ with certain parameter $\boldsymbol{\theta}$. The rationale of this systematic tuning stems from that the relevant quality of B_j decays when time of current measurement moves away from the time t_m of the primary endpoint. As a result, this tuning procedure has fewer parameters to be tuned. Various $\psi(\cdot)$ functions can be specified according to some preliminary knowledge of the underlying longitudinal relationship between time and the relative quality of outcomes. Some examples of $\psi(\cdot)$ are given as follows, and their time-course relevance to the endpoint are shown in Figure 4.3:

1. (Linear) $\psi(\tilde{t}_j \mid \theta) = (1 - \theta) + \theta\tilde{t}_j$, $\theta \in [0, 1]$.

2. (Exponential) $\psi(\tilde{t}_j \mid \boldsymbol{\theta}) = \theta_1 e^{\theta_2 \tilde{t}_j}$, $\theta_1 \in (0, 1], \theta_1 e^{\theta_2} = 1$.

3. (Polynomial) $\psi(\tilde{t}_j \mid \boldsymbol{\theta}) = \theta_1 \tilde{t}_j^2 + \theta_2 \tilde{t}_j + \theta_3$, $\theta_3 \in [0, 1], \theta_1 + \theta_2 + \theta_3 = 1$.

To make a choice between different $\psi(\cdot)$, cross validation may be employed according to the maximization of the estimated value function, defined as $\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))B/P(A \mid \boldsymbol{X})]/\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))/P(A \mid \boldsymbol{X})]$ (*Murphy et al.*, 2001), where $\mathbb{E}_n^*$

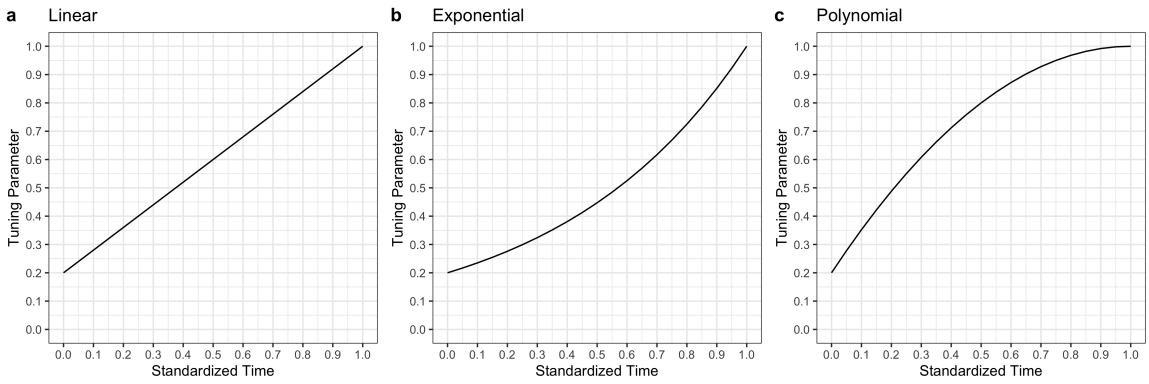

Figure 4.3: Change of tuning parameter λ_j according to standardized time $\tilde{t}_j$.

is the empirical average value and $D(\boldsymbol{X})$ is the assigned treatment based on the estimated ITR.

4.4 Simulation Experiment

4.4.1 Simulation Setting

This section reports one simulation experiment to illustrate the finite sample performance of scaled-tuning LS-learning in ITR derivation using longitudinal trial data with missing information. We randomly generated a 10-dimensional feature vector $\boldsymbol{X} = (X_1, \ldots, X_{10})^\intercal \in \mathbb{R}^{10}$ for each subject, with each feature $X_\nu \sim U(0,1), \nu = 1, \ldots, 10$ and $Corr(X_\nu, X_{\nu'}) = 0.2, \nu \neq \nu'$. Treatment $A \in \{-1, 1\}$ was randomly assigned with equal probability to each subject with $P(A = 1 \mid \boldsymbol{X}) = P(A = -1 \mid \boldsymbol{X}) = 0.5$. Subjects were measured at $m = 5$ follow-up times in addition to a baseline visit, with the scaled times as $t_j = j/m$ and the standardized time $\tilde{t}_j = t_j/t_m, j = 1, \ldots, m$. A set of cubic spline plus basis function $\boldsymbol{s}(t_j)$ was specified with three knots at $k_1 = 0.4, k_2 = 0.6$ and $k_3 = 0.8$, resulting in seven basis functions $(1, t_j, t_j^2, t_j^3, (t_j - k_1)_+^3, (t_j - k_2)_+^3, (t_j - k_3)_+^3)^\intercal$ at visit j to represent the time trajectory. The underlying true decision function f was specified as $f(\boldsymbol{X}) = 1 + X_1 - \log(X_2 + 1) + 2X_3^3 - \exp(X_4)$. The outcomes measured on subject i at time t_j was generated by an equation: $B_{ij} =$

$0.01+0.02X_{i,1}+\boldsymbol{s}(t_j)^\intercal\boldsymbol{\beta}_3+\{\boldsymbol{s}(t_j)^\intercal\boldsymbol{\beta}_4\}\{0.1(0.4X_{i,5}+0.6X_{i,6}-X_{i,7})\}+3A_if(\boldsymbol{X}_i)+\gamma_i+\epsilon_{ij}$, where $\boldsymbol{\beta}_3=\boldsymbol{\beta}_4=(3,0.5,0.5,-3.5,-2,-2,-0.1)^\intercal$, random intercepts $\gamma_i \overset{i.i.d}{\sim} N(0,0.5)$ and random errors $\boldsymbol{\epsilon}_i \overset{i.i.d}{\sim} MVN(\mathbf{0},\boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}(\rho)$ being an AR(1) correlation matrix and correlation coefficient $\rho=0.5$. The missing outcome indicator R_{ij} for each outcome B_{ij} was generated independently according to a Bernoulli distribution with probability (a) $P(R_{i1}=1)=0.01$ for the first measurement B_{i1} and (b) logit $P(R_{ij}=1)=1+2\tilde{t}_j-0.005B_{ij-1}$ for $j=2,\ldots,M$. The outcome B_{ij-1} at the adjacent visit $j-1$ was included to represent the not missing at random (NMAR) mechanism since B_{ij-1} may not be observed. Based on the missing patterns, subjects can be stratified into subgroups $\mathcal{S}_j, j=1,\ldots,m$ according to the available last observations. Although subjects were completely randomized to one of the two treatment groups, balancing may be broken due to missing data. Therefore, propensity scores of treatment assignment for each subject was estimated by a logistic regression model based on all the majors effects of the covariates $\boldsymbol{X}$. The total sample size for the training data $\mathcal{S}=\mathcal{S}_1\cup\cdots\cup\mathcal{S}_m$ was set as $n=1000$. In addition, an independent testing dataset of sample size 5000 with the true labels determined by $f(\boldsymbol{X})>0$ or not was generated as an external validation data. An ITR was derived using the LS-learning method under linear kernel. Scaled tuning was performed via an exponential function $\psi(\tilde{t}_j\mid\boldsymbol{\theta})=\theta_1e^{\theta_2\tilde{t}_j}$ with $\theta_2=\log(1/\theta_1)$ to characterize the relevant time course of the outcomes to the endpoint benefit measured at visit m. We set two tuning parameter pools $C\in\{1/128,1/64,\ldots,1/2,1\}$ and $\theta_1\in\{0.1,0.2,\ldots,0.9,1.0\}$ respectively, for tuning parameter selection. The GAM regression with B-spline basis functions (20 basis functions for each variable) was employed to estimate mean prediction function $g_j(\boldsymbol{X})$ and $h_j(\boldsymbol{s}(t_j))$ for parameter tuning. Simulation was repeated for 200 times to draw summary statistics.

4.4.2 Simulation Results

We compared five methods for their prediction accuracy on both training and validation datasets. They are, (M1) OWL on the whole training data $\mathcal{S}$ with no missing data. The endpoint B_m is fully observed and used for ITR derivation using the standard OWL. This serves as the gold standard, termed as the super oracle method, for the purpose of comparison. Note that the endpoint benefit B_m is often not available in practice due to missing data. (M2) LS-learning with scaled tuning on the whole training data $\mathcal{S}$ with missing data. It incorporates auxiliary benefits $B_j, j = 1, \ldots, m-1$ prior to the endpoint B_m using subject stratification. We derive ITR using LS-learning that takes into consideration the relevance of time course outcomes. (M3) OWL on the whole training data $\mathcal{S}$ with missing imputation under LOCF. It differs from M2 by ignoring the incorporation of time course outcomes. (M4) OWL on the whole training data with missing data, where missing information is imputed by multiple imputation via the R package MICE (*van Buuren and Groothuis-Oudshoorn*, 2011). The imputed endpoint B_m is used for ITR derivation via the standard OWL. Five imputed datasets were created using all the available data of $\boldsymbol{X}, A$ and observed benefits from which the average statistics of OWL performances are calculated. (M5) OWL on the training data of available endpoint benefits from those who are the completers of the clinical trial. All the methods based on the standard OWL, including M1, M3, M4, and M5, use 5-fold cross-validation for tuning parameter selection as proposed in *Zhao et al.* (2012). Table 4.2 lists the average prediction accuracy with the standard deviations on the training and validation datasets. It is shown that as expected, the super oracle M1 always gives the highest prediction accuracy since it uses the complete information for ITR derivation. However, M1 is not available in practice in the presence of missing data. The LS-learning method M2, which is developed with missing data, gives a slightly lower prediction accuracy compared to M1. Clearly, the prediction accuracy given by M3-M5 are much lower

Table 4.2: Average prediction accuracy (mean (sd)) on both training and validation datasets among the five methods to derive ITR.

Method	Prediction Accuracy (Training)	Prediction Accuracy (Validating)
M1	0.841 (0.034)	0.841 (0.033)
M2	0.815 (0.044)	0.813 (0.043)
M3	0.799 (0.048)	0.798 (0.047)
M4	0.743 (0.071)	0.741 (0.070)
M5	0.558 (0.096)	0.557 (0.097)

Table 4.3: Estimated value functions evaluated as B_1 to B_5 on the training data $\mathcal{S}$ for five methods of ITR derivation.

method	$\hat{V}(B_1)$	$\hat{V}(B_2)$	$\hat{V}(B_3)$	$\hat{V}(B_4)$	$\hat{V}(B_5)$
M1	10.07 (0.57)	10.07 (0.60)	9.70 (0.60)	8.78 (0.64)	6.90 (0.62)
M2	10.02 (0.58)	10.02 (0.60)	9.65 (0.60)	8.73 (0.66)	6.85 (0.63)
M3	9.98 (0.58)	9.99 (0.60)	9.60 (0.60)	8.69 (0.65)	6.82 (0.64)
M4	9.82 (0.60)	9.83 (0.63)	9.43 (0.62)	8.56 (0.66)	6.63 (0.66)
M5	9.24 (0.60)	9.22 (0.68)	8.87 (0.66)	7.95 (0.68)	6.12 (0.68)

than that of M2.

We calculated the estimated value functions evaluated at the five outcomes $\hat{V}(B_j)$ on the whole training data $\mathcal{S}$. These estimates are listed in Table 4.3. It is shown that the super oracle M1 always gives the highest estimated value function as expected. The LS-learning with scaled tuning method M2 gives higher estimated value functions than the other three OWL methods M3, M4, and M5. In LS-learning, the mean (sd) of the selected tuning parameter θ_1 in the exponential function equals 0.800 (0.270), resulting in the mean tuning parameters $\lambda_1 = 0.837, \lambda_2 = 0.875, \lambda_3 = 0.915, \lambda_4 = 0.956$ for B_1 to B_4, respectively. It demonstrates a decreasing trend for the relevance of outcomes from the endpoint to the first visit time. But the relevance remains strong since $\lambda_1 = 0.875$ is a relatively large weight of overlap for B_1 with B_5.

To compare computation time between LS-learning with greedy tuning and LS-learning with scaled tuning is of interest. Scaled tuning is proposed to save computation and costs for LS-learning by reducing the number of tuning parameters. In this simulation, should the greedy tuning be adopted, we need to tune a total of four λ

values ($\lambda_j, j = 1, \ldots, 4$), which is reduced to one by scaled tuning. The latter took, on average, 6,492.96 seconds, while the greedy tuning is estimated as 9,506,343 seconds (i.e. 110.03 days) under four sequence of grids $\lambda_j \in \{0, 0.1, \ldots, 1.0\}, j = 1, \ldots, 4$. Obviously, the computation efficiency of the scaled tuning method makes LS-learning practically feasible.

4.5 Derivation of ITR for Longitudinal Calcium Trial

We apply the LS-learning approach with the scaled-tuning scheme to analyze the longitudinal calcium supplementation trial. The central objective of our analysis is to derive an ITR that can guide pregnant women to take calcium supplementation in order to minimize persistent lead exposure for their children at age of 3 years old. The PBC values are measured repeatedly at a total of seven different follow-up times at month 3, 6, 12, 18, 24, 30, and 36, denoted as $B_1, \ldots, B_7$ respectively, where B_7 is designed as the primary endpoint of interest. It is reasonable to assume that the quality of these longitudinal outcomes increases when time of a PBC measurement approaches to month 36. A total of 376 mother-child pairs with no missing data at the baseline visit are used for our ITR derivation, with 190 mothers taking calcium supplementation and 186 mothers taking placebo. Stratification of mother-child pairs into strata $\mathcal{S}_1, \ldots, \mathcal{S}_7$ based on the primary endpoint for the completers and surrogate endpoints for the noncompleters are illustrated in Figure 4.1(b). Section 4.2 presents more details of the study design and data of the trial.

Since children on the two treatment arms are not fully balanced due to missing data, propensity scores are estimated by a logistic regression model with all the major effects of the predictors. Similar to the analysis performed in Section 3.5, after weighting the PBC values by the estimated propensity scores, we evaluate the balance of the propensity scores in Figure 4.4(a) and more importantly the balance of individual predictor's distribution using the absolute standardized mean difference

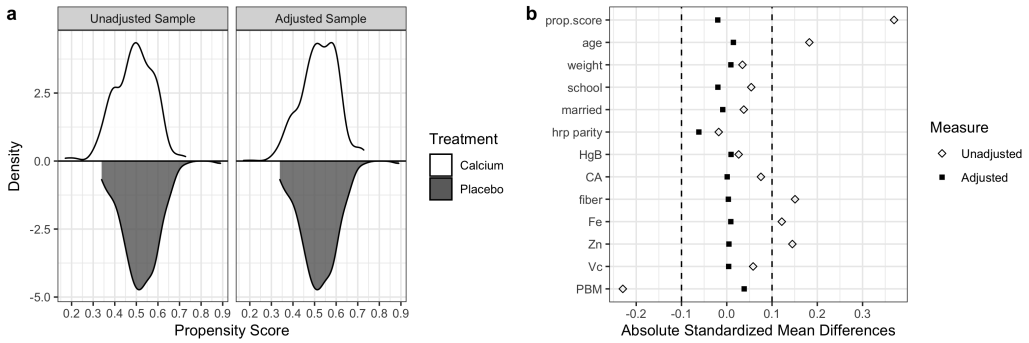

Figure 4.4: (a) Distributional balance of the propensity scores before and after weighting. (b) Balance of individual predictor before and after weighting adjusted by ASMD under threshold 0.1, indicated by the dashed lines.

(ASMD) in Figure 4.4(b). Figure 4.4(b) shows clearly that the balance of predictor's distribution is satisfactorily reached, as for all individual predictors their ASMDs (denoted by solid dots) have met within a threshold of 0.1 (*Austin*, 2009).

We now derive ITR for the daily intake of calcium supplementation. We perform: (i) a standard OWL on the stratum of completers, $\mathcal{S}_7$ with PBC B_7 (or PBC36), and the resulting decision rule is denoted as $\hat{f}_1$; (ii) a standard OWL on the whole training data $\mathcal{S}$ with missing data imputed by the method of multiple imputation via the R package MICE (*van Buuren and Groothuis-Oudshoorn*, 2011), resulting in the estimated decision rule $\hat{f}_2$; (iii) a standard OWL on the whole training data $\mathcal{S}$ with missing PBC36 replaced by the last observed PBC under the so-called strategy of last observation carried forward (LOCF) (*Molnar et al.*, 2008). In effect, LOCF implies that the last observed PBC is used as endpoint outcome for ITR derivation. Clearly, this approach ignores the differential relevance of longitudinal outcomes in the reference to the primary endpoint PBC36 measured at month 36. From the tuning parameter point of view, this is equivalent to setting naively $\lambda_j = 1, j = 1, \ldots, 6$. The estimated decision rule is denoted as $\hat{f}_3$; (iv) our LS-learning with the scaled tuning scheme on the whole training data $\mathcal{S}$, which takes into account differences among longitudinal outcomes $B_j, j = 1, \ldots, 6$ in terms of their relevance to PBC36, and this leads to an estimated decision rule $\hat{f}_4$. In LS-learning, we use

Table 4.4: Estimated intercepts and coefficients of the predictors in the estimated ITRs.

	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$
intercept	0.40	0.08	0.08	0.18
age	0.45	0.37	0.35	0.51
weight	0.10	0.02	0.03	0.03
school	0.01	-0.04	-0.02	-0.03
married	0.20	0.17	0.14	0.19
parity	-0.55	0.01	0.12	-0.16
HgB	0.12	0.13	0.13	0.10
Ca	-0.42	-0.28	-0.28	-0.46
fiber	0.34	0.23	0.24	0.26
Fe	-0.11	0.05	0.03	0.08
Zn	0.26	0.29	0.32	0.53
VC	0.06	-0.15	-0.19	-0.04
PBM	-0.40	-0.43	-0.43	-0.47

the linear kernel in the optimization function (4.1) to train ITR with the utility of the exponential scaling function for tuning (the middle panel of Figure 4.3). We set $C \in \{1/128, 1/64, \ldots, 1/2, 1\}$ and $\theta_1 \in \{0.1, 0.2, \ldots, 0.9, 1.0\}$ for the tuning parameter selection. The loading coefficients of the predictors in the above four estimated ITRs are listed in Table 4.4.

According to the exponential scaling procedure, we obtain a tuning parameter $\theta_1 = 0.1$, which generates $\lambda_1 = 0.121, \lambda_2 = 0.147, \lambda_3 = 0.215, \lambda_4 = 0.316, \lambda_5 = 0.464$ and $\lambda_6 = 0.681$ for $\hat{f}_4$. In this LS-learning, we use two tuning parameters, C and θ_1, each having 8 and 10 different values, respectively, resulting in a total of 80 different pairs of C and θ_1. In contrast, if we use a greedy tuning strategy under $\lambda_j \in \{0, 0.1, \ldots, 1.0\}, j = 1, \ldots, 6$, we would have to deal with a total of $8 \times 11^6 = 14,172,488$ pairs in tuning. As a result, our scaling procedure has saved 177,156 folds of computational runs.

Applying each of the four estimated decision rules above, we would allocate a pregnant woman to take calcium supplement if $\hat{f} > 0$; otherwise, not to take the supplement. Among 376 mothers, $\hat{f}_1, \hat{f}_2, \hat{f}_3$ and $\hat{f}_4$ would assign 233 (62.0%), 243 (64.6%),

Table 4.5: Treatment assignment comparison between complete randomization and $\hat{f}_4$ estimated by LS-learning.

		$\hat{f}_4$		
		placebo	calcium	total
randomization	placebo	87	90	177
	calcium	59	131	190
	total	146	221	367

254 (67.6%) and 230 (61.2%) women to take calcium supplement, respectively, all higher than 50% given in the randomized assignment. In particular, comparing between complete randomization and the LS-learning $\hat{f}_4$, as shown in Table 4.5, we see significant differences in the reallocation of calcium supplement. We conduct the McNemar test for the hypothesis of homogeneous allocation distribution, and obtain p-value of 0.0015, indicating that there exists a significant discrepancy between the two allocation rules, and in other words, these different treatment assignments do not occur by accident.

In addition, we calculate the SSE values to assess the quality of prediction by $\hat{f}_1, \hat{f}_2, \hat{f}_3$ and $\hat{f}_4$. They are 1117.2, 1113.5, 1114.1 and 923.3, respectively. This suggests that the LS-learning $\hat{f}_4$ achieves the smallest SSE or the best prediction of the benefit values. To compare the performances of $\hat{f}_1, \hat{f}_2, \hat{f}_3$ and $\hat{f}_4$ in minimizing continuous lead exposure over different time points, we calculate the estimated value functions $\hat{V}(B_j)$ of $B_j, j = 1, \ldots, 7$ using $\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))B/P(A \mid \boldsymbol{X})]/\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))/P(A \mid \boldsymbol{X})]$ (*Murphy et al.*, 2001). The estimated value functions for longitudinal outcomes at individual visit times as well as their averages are summarized in Table 4.6. Clearly, the LS-learning $\hat{f}_4$ gives the highest average estimated value function, producing the lowest continuous lead exposure in comparison to the other three decision rules $\hat{f}_1, \hat{f}_2$ or $\hat{f}_3$. This indicates that the utility of the longitudinal auxiliary data via surrogate outcomes $B_j, j = 1, \ldots, 6$ has arguably improved the estimated ITR to maximally benefit children in their growth and development. This improve-

Table 4.6: Estimated value functions at each visit and the average based on the estimated ITR $\hat{f}_1, \hat{f}_2, \hat{f}_3$ and $\hat{f}_4$ derived in the calcium supplementation trial.

ITR	$\hat{V}(B_1)$	$\hat{V}(B_2)$	$\hat{V}(B_3)$	$\hat{V}(B_4)$	$\hat{V}(B_5)$	$\hat{V}(B_6)$	$\hat{V}(B_7)$	Average
$\hat{f}_1$	9.1	10.6	9.4	6.3	8.1	8.1	8.3	8.56
$\hat{f}_2$	9.6	10.2	9.4	6.3	8.5	8.1	8.3	8.63
$\hat{f}_3$	9.6	9.6	9.4	6.8	8.4	8.1	8.3	8.60
$\hat{f}_4$	9.6	10.6	9.4	6.3	8.5	8.1	8.3	8.69

ment is rooted in a successful use on the ordering of time-course relevance among longitudinal outcomes in the multi-view extension of OWL with multiple training datasets.

We also compare individual predictor's distributions between the two resulting treatment groups by the LS-learning decision function $\hat{f}_4$; see Table 4.7, through which we identify those predictors that are significantly different between the two groups. We found that mothers assigned by $\hat{f}_4$ to take calcium supplementation have a significantly older age, higher weight, longer years in school, higher proportion of marriage, higher dietary intake of fiber, Fe, Zn and vitamin C, and lower maternal baseline blood lead concentration (PBM) (see Table 4.7). It is interesting to note that mothers assigned to the calcium supplement group by $\hat{f}_4$ do not have a significant difference in their dietary calcium intake from food compared to those assigned to the placebo group (11.4 (5.2) versus 10.4 (5.2)). This suggests that maternal dietary intake of calcium on baseline may not influence children's persistent lead exposure at age 3. This finding is of clinical value, as some pregnant women refuse to take calcium supplementation under their misbelief of being able to obtain adequate amount of calcium from food. In addition, no significant differences were identified for the total number of pregnancies and the maternal baseline HgB concentration between the two resulting allocation groups by $\hat{f}_4$.

Table 4.7: Summary statistics of the predictors based on treatment allocation according to $\hat{f}_4$. Mean (sd) and percentage values are shown, where p-values are obtained from Wilcoxon rank-sum test and chi-aquare test for numeric and categorical variables, respectively.

Predictor	Calcium	Placebo	p-value
age	28.3 (5.2)	23.4 (4.7)	4.36×10^{-17}
weight	62.9 (10.9)	60.1 (9.7)	9.56×10^{-3}
school	11.2 (3.1)	10.1 (2.4)	2.14×10^{-4}
married	0.735	0.610	1.49×10^{-2}
parity	2.1 (1.0)	2.0 (1.0)	3.30×10^{-1}
HgB	13.3 (0.8)	13.0 (1.2)	8.84×10^{-2}
Ca	11.4 (5.2)	10.4 (5.2)	8.15×10^{-2}
fiber	26.4 (9.2)	19.2 (8.3)	1.33×10^{-14}
Fe	14.9 (5.6)	10.8 (5.0)	1.76×10^{-12}
Zn	10.5 (3.4)	7.9 (3.2)	7.44×10^{-13}
VC	193.6 (84.7)	146.9 (81.6)	1.43×10^{-8}
PBM	4.0 (2.3)	6.4 (3.9)	2.64×10^{-11}

4.6 Concluding Remarks

In this chapter, we utilize a newly proposed LS-learning method to establish ITRs in that we integrate longitudinal data sources with time-varying relevance to the primary clinical endpoint. Our method provides a useful extension of the existing one-dimensional cross-sectional OWL method by allowing temporally correlated outcomes in the search of optimal ITR. We introduce additional tuning parameters to synergize different training data sources with varying degrees of relevance to the primary data source. The proposed method is particularly appealing in real clinical studies where repeated measurements are often collected with missing data that would lead to data attrition and loss of statistical power if no appropriate strategy is invoked to handle this issue. Meanwhile, we propose and test a scaled tuning procedure in LS-learning to greatly reduce computational burden in the tuning parameter selection.

We apply LS-learning to derive an ITR for pregnant women in their daily intake of calcium supplementation to lower persistent lead exposure by their children at age 3 years old. The major difficulty in analyzing this clinical trial data lies in the fact that

there are a substantial amount of missing values of PBC over the three-year period of the study. The LS-learning stratifies subjects into subgroups of outcomes based on respective missing patterns in a similar way considered in the pattern mixture model proposed by *Little* (2008). It is shown that the new ITR derived by LS-learning would lead to a larger reduction of persistent lead exposure to children in comparison to other ITRs given by standard OWL, if the new one were implemented for the whole target population of pregnant women. In addition to the real data application, a comprehensive simulation experiment in Section 4.4 shows that LS-learning outperforms the standard OWL in ITR estimation according to the prediction accuracy as well as the estimated value functions. Meanwhile, we have identified several baseline predictors that may play important roles in allocating calcium supplementation, such as dietary intake of fiber, Fe, Zn and vitamin C.

One limitation of our data analysis is the use of linear kernel for ITR derivation because of its ease in illustrating the estimated decision rules. Nonlinear-kernels can also be applied to derive more flexible decision rules, which may be further explored as a future extension. Another challenge pertains to the imbalance of treatment assignment in the cleaned data caused mostly by missing baseline predictors. We have adopted the means of propensity score via inverse probability weighting (IPW) to adjust the treatment allocation bias. Other methods, like augmented inverse probability weighting (AIPW) (*Glynn and Quinn*, 2010) and overlap weighting (*Thomas et al.*, 2020), are worth a further exploration to overcome the treatment assignment bias in the ITR analysis of calcium supplementation clinical trial.

CHAPTER V

Summary and Future Work

Focusing on individualized treatment rules (ITRs) in the filed of personalized healthcare, this dissertation has delivered improvements and extensions of ITR methodologies, with specific applications in precision medicine and precision nutrition. Various methods have been developed to derive ITR that allows to tailor treatments for patients with heterogeneous responses, and consequently maximizes their therapeutic effects. However, many methodological challenges still remain to be addressed for better ITR. Some of the ITR improvements considered in this dissertation include (i) to improve and validate existing ITRs through the use of potential new candidate biomarkers, (ii) to establish ITRs via a multi-view approach with multiple benefit outcomes, and (iii) to derive ITRs in the presence of missing information in training data. These tasks are not trivial analytically, while are of critical importance in real clinical practice.

The objective of validating existing ITRs using candidate biomarkers has led to the development of the net benefit index (NBI) method in Chapter II. We proposed a new statistical framework, NBI, to quantify added values of candidate biomarkers and select those significantly important ones to be included in an existing ITR. This selection is done by a rigorous NBI-based test. Another difficult task in ITR derivation is to establish ITR in the presence of multiple benefit outcomes, which are typically

of various quality and clinical relevance with missing data. These complications are encountered in the ITR derivation for precision nutrition, where patients have been observed for multiple benefit responses in a longitudinal clinical trial. Driven by these complicating factors, we developed two methods, termed synergistic self-learning (SS-learning) and longitudinal self-learning (LS-learning), respectively, in Chapter III and Chapter IV. SS-learning and LS-learning are two self-iterative procedures that allow the integration of multiple benefit outcomes into ITR derivation, which extends the existing one-dimensional method for of ITR derivation.

The three methods proposed in this dissertation can be further improved as have been discussed at the end of the respective chapters. For future research, the effort should be made on borrowing the strengths of the proposed methods and pairing them with new methodologies in the fast growing literature in machine learning for further generalization of ITR methodology. Here, we conclude this dissertation by pointing out several potential future research directions on ITR derivation for personalized healthcare.

One promising direction of future research is to extend the derivation of ITR from trials with categorized treatment levels to clinical studies of continuous dosage amounts. One limitation of the proposed NBI method, SS-learning, and LS-learning is that they can only deal with the allocation of binary treatment arms, for example, new treatment and standard treatment or placebo. Such an ITR estimation problem with categorical treatment levels may actually be formulated as a classification problem, which is solved by certain machine learning methods like support vector machine (SVM). One of such example is outcome weighted learning (OWL) reported by *Zhao et al.* (2012). However, in real clinical practice, medications and therapies may be allocated to patients with some continuous dosage amounts. Therefore, it seems of great interest to estimate the optimal individualized dosage of medicine in real clinical studies. Methods have already been developed for dose finding in ITR derivation,

like *Chen et al.* (2016), in which the authors propose an OWL-based method with a nonconvex loss function that can be solved by an algorithm based on a difference of convex functions. Kernel-based and Bayesian methodologies are also emerging in optimal individualized dose finding; see for examples *Zhu et al.* (2020) and *Liu and Johnson* (2016). Future research is needed to develop robust and clinical relevant methods of continuous dose finding for ITR derivation.

In the advent of electronic health records (EHR) data, another promising future research direction is to establish ITR using big data from observational studies, such as those EHR databases. Different from randomized trials, subjects' records in EHR databases do not receive random treatment assignments. This kind of non-experimental study suffers from selection bias and unmeasured confounding, which challenge the establishment of inference for causality. In this case, some propensity score methods are inevitable to adjust high-dimensional confounding and selection biases. One recent publication (*Wu et al.*, 2020) proposes a machine learning approach based on matching (M-learning) to deal with the problem of ITR derivation using EHR data. This newly proposed M-learning method employs matching instead of inverse probability weighting (IPW), which is the approach that is commonly used in many existing methods like OWL for ITR estimation, to better evaluate individual responses to different treatments and alleviate confounding associated with EHR data. The proposed matching-based value function can serve as a unified framework that can accommodate the use of other types of outcomes, such as continuous, ordinal, and discrete outcomes, in ITR derivation. This also provides another promising direction to extend the current ITR derivation methods to deal with various benefit outcomes types. Along this line of future research, other propensity score matching methods or weighting schemes can also be examined in the setting of ITR estimation under the M-learning framework using EHR data.

Last but not least, we plan to develop online ITR estimation methods using

streaming data, with the objective that we want to update ITR by some online estimation methods when training data become available sequentially over time. An online ITR avoids re-estimating ITR using the entire training dataset that grows sequentially. With the quick growth of EHR data, online ITR estimation methods that can deal with growing large EHR databases are appealing. It allows us to establish robust and reproducible decision functions in a reasonable amount of time. Streaming classification is a topic that has been studied in machine learning (*Mu et al.*, 2017; *Abdulsalam et al.*, 2010; *Street and Kim*, 2001; *Nair et al.*, 2018). Such online classification methods are able to handle intermittent arrival of labeled records, and to adjust parameters in respond to changing class boundaries in data streams. We plan to follow these online classification methods for ITR derivation using streaming EHR data sources.

APPENDICES

APPENDIX A

Outcome Weighted Learning (OWL)

Denote the space of the observed data as $(\boldsymbol{X}, A, B)$. The distribution of $(\boldsymbol{X}, A, B)$ is denoted as P and the expectation of B with respect to P is denoted as E. Given a specific decision rule D, the distribution of $(\boldsymbol{X}, A, B)$ where $A = D(\boldsymbol{X})$ is denoted as P^D, and the expectation of B with respect to P^D is denoted as E^D. The expected clinical benefit under the given D can be calculated as the value function

$$V(D) = E^D(B) = \int B dP^D = \int B \frac{dP^D}{dP} dP = E\left\{ \frac{I(A = D(\boldsymbol{X}))}{P(A_i \mid \boldsymbol{X}_i)} B \right\}. \tag{A.1}$$

The optimal decision rule will be D^* such that

$$D^* \in \arg\max_D E\left\{ \frac{I(A = D(\boldsymbol{X}))}{P(A_i \mid \boldsymbol{X}_i)} B \right\}, \tag{A.2}$$

which is equivalent to

$$D^* \in \arg\min_D E\left\{ \frac{I(A \neq D(\boldsymbol{X}))}{P(A_i \mid \boldsymbol{X}_i)} B \right\}. \tag{A.3}$$

The term $\frac{I(A \neq D(\boldsymbol{X}))}{P(A_i|\boldsymbol{X}_i)} B$ is actually a weighted classification error. Therefore, OWL is a weighted classification problem. With a set of $i.i.d$ observations $(\boldsymbol{X}_i, A_i, B_i)$, we

can approximate the optimization problem in (A.3) by the empirical value

$$D^* \in \arg\min_D \frac{1}{n} \sum_{i=1}^{n} \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)} I\{A_i \neq D(\boldsymbol{X}_i)\}. \tag{A.4}$$

Since $D(\boldsymbol{X})$ can always be represented as $\text{sign}(f(\boldsymbol{X}))$ for some decision function f, where

$$D(\boldsymbol{X}) = \begin{cases} 1, & f(\boldsymbol{X}) > 0, \\ -1, & f(\boldsymbol{X}) < 0, \end{cases}$$

the optimization problem in (A.4) is equivalent to:

$$f^* \in \arg\min_f \frac{1}{n} \sum_{i=1}^{n} \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)} I\{A_i \neq \text{sign}(f(\boldsymbol{X}_i))\} \tag{A.5}$$

We can first obtain the optimal decision function f^* based on the optimization of (A.5), and then set $D^*(\boldsymbol{X}) = \text{sign}(f^*(\boldsymbol{X}))$ to get the optimal decision rule D^*.

APPENDIX B

Support Vector Machine (SVM) solution to OWL

The optimization problem (A.5) is a weighted summation of 0-1 loss, which is neither convex nor continuous. It makes the problem difficult to be solved directly. Therefore, OWL uses a convex surrogate hinge loss $x^+ = \max(0, x)$, which is commonly used in SVM, to replace the 0-1 loss. In order to further penalize the complexity of the decision function f to avoid overfitting, OWL adds a l_2 penalty into the optimization problem. The final function OWL aims to minimize is

$$\frac{1}{n}\sum_{i=1}^{n}\frac{B_i}{P(A_i \mid \boldsymbol{X}_i)}(1 - A_i f(\boldsymbol{X}_i))^+ + \lambda_n ||f||_2, \tag{B.1}$$

where λ_n is the regulization parameter. This optimization problem can be solved using the technique of SVM.

If we assume that the decision rule f is a linear function $f(\boldsymbol{X}) = \beta_0 + \boldsymbol{X}\boldsymbol{\beta}$, the optimization problem of OWL can be solved as follows by introducing the slack

variable $\xi_i = (1 - A_i f(\boldsymbol{X}_i))^+$,

$$\begin{aligned} \text{minimize} \quad & \frac{1}{n}\sum_{i=1}^{n} \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)}\xi_i + \lambda_n ||\beta||_2 \\ \text{subject to} \quad & A_i(X_i^T\beta + \beta_0) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad i = 1, ..., n. \end{aligned} \tag{B.2}$$

Let $\kappa = \frac{1}{2n\lambda_n}$, the optimization problem is transformed to

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}||\beta||_2 + \kappa \sum_{i=1}^{n} \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)}\xi_i \\ \text{subject to} \quad & A_i(X_i^T\beta + \beta_0) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad i = 1, ..., n. \end{aligned} \tag{B.3}$$

Solve (B.3) by introducing the Lagrange Multipliers, we come to the dual problem

$$\begin{aligned} \underset{\alpha}{\text{maximize}} \quad & \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j A_i A_j X_i \cdot X_j \\ \text{subject to} \quad & 0 \leq \alpha_i \leq \kappa \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)} \\ & \sum_{i=1}^{n} \alpha_i A_i = 0, \end{aligned} \tag{B.4}$$

where α_i is the Lagrange Multiplier. If we assume a nonlinear decision rule f, the optimization problem of OWL can be solved using the kernal function k,

$$\begin{aligned} \underset{\alpha}{\text{maximize}} \quad & \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j A_i A_j k(X_i, X_j) \\ \text{subject to} \quad & 0 \leq \alpha_i \leq \kappa \frac{B_i}{P(A_i \mid \boldsymbol{X}_i)} \\ & \sum_{i=1}^{n} \alpha_i A_i = 0. \end{aligned} \tag{B.5}$$

APPENDIX C

Multiple-variable-based ITR Derivation when $\boldsymbol{X}_e = Null$ for Chapter II

This simulation concerns a setting where $\boldsymbol{X}_e = Null$. We have multiple signal and noise candidate biomarkers, $X_j \sim U(0,1), j = 1, \ldots, 10$, in which only X_1 and X_2 are signal biomarkers involved in the optimal ITR. The correlation structure of the variables is that $\text{Corr}(X_1, X_3) = \text{Corr}(X_2, X_4) = 0.5$, and $\text{Corr}(X_s, X_t) = 0.2, s, t \in \{5, \ldots, 110\}, s \neq t$. B is generated from a normal distribution with mean $\mu = 0.5 + X_1 + 2.0Af(X)$, where $f(\boldsymbol{X})$ is given as follows:

7) (Linear) $f(\boldsymbol{X}) = 0.5(1 + 2X_1 - 4X_2)$;

8) (Binary) $f(\boldsymbol{X}) = 6\{I(X_1 > 0.29 \cap X_2 < 0.71) - 0.5\}$;

9) (Nonlinear) $f(\boldsymbol{X}) = (X_1 - 0.1)^+ - (X_2 - 0.22)^+$.

Summary statistics of variable selection based on the proposed NBI test method, SAS and riskRFE are included in Table C.1.

Table C.1: Size, TDR, MCC, and CCR for variable selection based on NBI test, SAS, and riskRFE in the multiple-variable-based decision rule evaluation when $\boldsymbol{X}_e = Null$.

NBI					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	800	1.799 (0.761)	0.915 (0.192)	0.797 (0.221)	0.871 (0.068)
	1000	1.895 (0.777)	0.918 (0.181)	0.826 (0.207)	0.880 (0.067)
	1200	1.934 (0.707)	0.917 (0.184)	0.844 (0.211)	0.881 (0.069)
binary	800	2.115 (0.617)	0.935 (0.151)	0.921 (0.149)	0.905 (0.072)
	1000	2.145 (0.590)	0.937 (0.150)	0.933 (0.142)	0.909 (0.070)
	1200	2.136 (0.467)	0.948 (0.134)	0.953 (0.120)	0.914 (0.061)
nonlinear	800	1.957 (0.840)	0.898 (0.216)	0.814 (0.243)	0.840 (0.087)
	1000	2.045 (0.914)	0.893 (0.209)	0.828 (0.232)	0.847 (0.080)
	1200	1.977 (0.771)	0.912 (0.195)	0.842 (0.224)	0.855 (0.081)
SAS					
scenario	n	size(sd)	TDR(sd)	MCC(sd)	CCR(sd)
linear	800	2.096 (0.295)	0.968 (0.098)	0.977 (0.070)	0.988 (0.007)
	1000	2.034 (0.192)	0.989 (0.061)	0.992 (0.044)	0.989 (0.006)
	1200	2.022 (0.147)	0.993 (0.049)	0.995 (0.035)	0.991 (0.005)
binary	800	2.502 (0.712)	0.852 (0.190)	0.891 (0.143)	0.825 (0.011)
	1000	2.316 (0.584)	0.904 (0.166)	0.930 (0.123)	0.828 (0.011)
	1200	2.208 (0.457)	0.934 (0.139)	0.953 (0.101)	0.826 (0.012)
nonlinear	800	3.228 (1.065)	0.688 (0.219)	0.760 (0.179)	0.963 (0.015)
	1000	2.840 (0.903)	0.769 (0.213)	0.827 (0.166)	0.964 (0.015)
	1200	2.672 (0.826)	0.810 (0.207)	0.859 (0.158)	0.967 (0.012)
riskRFE					
scenario	n	size(sd)	TDR(sd)	MCC(sd)	CCR(sd)
linear	800	2.686 (0.955)	0.756 (0.229)	0.774 (0.217)	0.925 (0.042)
	1000	2.443 (0.829)	0.813 (0.228)	0.815 (0.227)	0.928 (0.045)
	1200	2.219 (0.676)	0.870 (0.196)	0.854 (0.198)	0.932 (0.035)
binary	800	3.312 (1.037)	0.646 (0.209)	0.714 (0.192)	0.859 (0.108)
	1000	2.985 (0.968)	0.710 (0.215)	0.763 (0.186)	0.883 (0.093)
	1200	2.638 (0.904)	0.783 (0.216)	0.810 (0.192)	0.885 (0.085)
nonlinear	800	2.969 (1.087)	0.672 (0.244)	0.691 (0.249)	0.842 (0.069)
	1000	2.645 (0.958)	0.741 (0.244)	0.746 (0.245)	0.845 (0.071)
	1200	2.401 (0.842)	0.789 (0.236)	0.773 (0.241)	0.853 (0.076)

APPENDIX D

Additional Simulation Experiments for Chapter II

Table D.1 lists the summary statistics of variable selection based on the proposed NBI method, SAS and riskRFE in the multiple-variable-based decision rule evaluation when $\boldsymbol{X}_e \neq Null$ and $n = 200$. Table D.2 lists the summary statistics of variable selection based on the proposed NBI test method, SAS and riskRFE in the multiple-variable-based decision rule evaluation when $\boldsymbol{X}_e = Null$ and NBI is calculated under the Gaussian kernel. In this simulation setting, the true decision rule f is set as $f(\boldsymbol{X}) = \exp(X_1^2) - \exp(X_2^2)$. Table D.3 is the comparison of prediction accuracy for the ITRs derived under NBI test and the standard OWL in the multiple-variable-based decision rule evaluation when $\boldsymbol{X}_e \neq Null$. Table D.4 lists the summary statistics of variable selection based on l_1-OWL in the multiple-variable-based decision rule evaluation when $\boldsymbol{X}_e = Null$.

We also report the simulation results (see Table D.5-D.7) for an additional single-variable-based nonlinear decision rule ($f(\boldsymbol{X}) = 1 + X_1 + X_2^3 - \exp(X_3)$) and an additional multiple-variable-based nonlinear decision rule ($f(\boldsymbol{X}) = 1.5\{1 + X_1 - \log(X_2 + 1) + 2X_3^3 - \exp(X_4)\}$). It is seen that NBI still gives good variable selection results in the single-variable-based decision rule evaluation, while in the multiple-variable-based decision rule evaluation, SAS outperforms NBI. It is interesting to note

Table D.1: Size, TDR, MCC, and CCR for variable selection based on NBI test, SAS and riskRFE in the multiple-variable-based decision rule evaluation when $n = 200$.

NBI				
scenario	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	1.321 (1.327)	0.549 (0.443)	0.414 (0.364)	0.707 (0.106)
binary	1.235 (1.186)	0.538 (0.447)	0.403 (0.371)	0.654 (0.081)
nonlinear	1.637 (1.685)	0.580 (0.431)	0.430 (0.372)	0.686 (0.108)
SAS				
scenario	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	6.819 (2.154)	0.323 (0.140)	0.316 (0.239)	0.918 (0.039)
binary	7.626 (1.984)	0.262 (0.104)	0.189 (0.236)	0.723 (0.023)
nonlinear	7.246 (2.046)	0.295 (0.120)	0.271 (0.226)	0.889 (0.043)
riskRFE				
scenario	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	5.325 (1.503)	0.330 (0.156)	0.296 (0.301)	0.747 (0.081)
binary	6.223 (1.418)	0.321 (0.102)	0.342 (0.204)	0.645 (0.100)
nonlinear	5.304 (1.570)	0.323 (0.156)	0.286 (0.291)	0.727 (0.081)

that in scenario $n = 1200$, SAS gives the perfect results with size=2.000, TDR=1.000 and MCC=1.000, with ZERO standard deviation. This does not seem to be an appropriate setting to reflect a real-world scenario. Such perfection may be resulted from too strong signal-to-noise ratio in the previously chosen nonlinear decision rule.

Table D.2: Size, TDR, MCC, and CCR for variable selection based on NBI test, SAS and riskRFE in the multiple-variable-based decision rule evaluation when SVM is performed under Gaussian kernel.

NBI					
scenario	n	size(sd)	TDR(sd)	MCC(sd)	CCR(sd)
nonlinear	800	2.125 (0.413)	0.958 (0.116)	0.966 (0.090)	0.878 (0.062)
	1000	2.175 (0.506)	0.949 (0.132)	0.962 (0.099)	0.884 (0.061)
	1200	2.150 (0.410)	0.951 (0.124)	0.963 (0.092)	0.890 (0.052)
SAS					
scenario	n	size(sd)	TDR(sd)	MCC(sd)	CCR(sd)
nonlinear	800	2.327 (0.604)	0.902 (0.168)	0.928 (0.125)	0.990 (0.007)
	1000	2.183 (0.426)	0.941 (0.131)	0.958 (0.095)	0.991 (0.006)
	1200	2.105 (0.341)	0.967 (0.104)	0.976 (0.075)	0.992 (0.005)
riskRFE					
scenario	n	size(sd)	TDR(sd)	MCC(sd)	CCR(sd)
nonlinear	800	2.320 (0.467)	0.783 (0.279)	0.776 (0.338)	0.811 (0.109)
	1000	2.200 (0.400)	0.798 (0.293)	0.777 (0.360)	0.819 (0.106)
	1200	2.168 (0.374)	0.816 (0.277)	0.795 (0.340)	0.832 (0.101)

Table D.3: Prediction accuracy for ITRs derived by NBI and the standard OWL in the multiple-variable-based decision rule evaluation.

scenario	n	CCR (sd) NBI	CCR (sd) OWL
linear	800	0.835 (0.081)	0.837 (0.037)
	1000	0.852 (0.076)	0.850 (0.039)
	1200	0.870 (0.070)	0.862 (0.035)
binary	800	0.765 (0.098)	0.649 (0.121)
	1000	0.786 (0.095)	0.667 (0.125)
	1200	0.805 (0.090)	0.666 (0.125)
nonlinear	800	0.818 (0.081)	0.821 (0.042)
	1000	0.832 (0.075)	0.830 (0.041)
	1200	0.847 (0.073)	0.841 (0.037)

Table D.4: Size, TDR, MCC and CCR in the multiple-variable-based decision rule evaluation based on l_1-OWL when $\boldsymbol{X}_e = Null$.

scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
linear	800	6.031 (3.358)	0.212 (0.197)	0.022 (0.288)	0.771 (0.161)
	1000	6.337 (3.203)	0.190 (0.159)	-0.013 (0.289)	0.785 (0.162)
	1200	6.744 (3.040)	0.190 (0.133)	-0.005 (0.271)	0.802 (0.156)
binary	800	5.950 (3.416)	0.187 (0.167)	-0.002 (0.276)	0.653 (0.144)
	1000	6.226 (3.163)	0.189 (0.159)	-0.012 (0.289)	0.665 (0.147)
	1200	6.757 (3.143)	0.199 (0.139)	0.005 (0.274)	0.679 (0.147)
nonlinear	800	5.887 (3.353)	0.196 (0.179)	0.003 (0.280)	0.731 (0.115)
	1000	6.427 (3.211)	0.199 (0.165)	-0.010 (0.290)	0.758 (0.110)
	1200	6.543 (2.968)	0.197 (0.144)	-0.004 (0.297)	0.769 (0.107)

Table D.5: Discovery rates for X_3 and X_4 in the single-variable-based decision rule evaluation for the additional nonlinear setting. (Discovery rate for X_4 equals 1-specificity.)

		$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
scenario	n	X_3	X_4	X_3	X_4	X_3	X_4	X_3	X_4
nonlinear	800	0.977	0.055	0.977	0.054	0.965	0.058	0.855	0.051
	1000	0.981	0.057	0.985	0.057	0.969	0.054	0.872	0.045
	1200	0.991	0.052	0.989	0.037	0.98	0.045	0.918	0.048

Table D.6: NBI values for X_3 and X_4 in the single-variable-based decision rule evaluation for the additional nonlinear decision rule.

		$\rho = 0.0$		$\rho = 0.2$	
scenario	n	X_3 mean (sd)	X_4 mean (sd)	X_3 mean (sd)	X_4 mean (sd)
nonlinear	800	1.443 (0.609)	-0.151 (0.633)	1.424 (0.604)	-0.153 (0.662)
	1000	1.405 (0.584)	-0.124 (0.613)	1.404 (0.572)	-0.112 (0.613)
	1200	1.421 (0.517)	-0.122 (0.595)	1.379 (0.520)	-0.168 (0.562)
		$\rho = 0.5$		$\rho = 0.8$	
scenario	n	X_3 mean (sd)	X_4 mean (sd)	X_3 mean (sd)	X_4 mean (sd)
nonlinear	800	1.266 (0.642)	-0.158 (0.646)	0.863 (0.646)	-0.143 (0.642)
	1000	1.233 (0.646)	-0.110 (0.600)	0.819 (0.649)	-0.139 (0.623)
	1200	1.221 (0.560)	-0.159 (0.595)	0.867 (0.571)	-0.123 (0.611)

Table D.7: Size, TDR, MCC, and CCR for variable selection based on NBI test, SAS and riskRFE in the multiple-variable-based decision rule evaluation for the additional nonlinear decision rule.

NBI					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
nonlinear	800	2.064 (0.588)	0.936 (0.161)	0.912 (0.179)	0.910 (0.007)
	1000	2.085 (0.506)	0.943 (0.149)	0.932 (0.156)	0.912 (0.007)
	1200	2.085 (0.465)	0.946 (0.141)	0.938 (0.148)	0.912 (0.007)
SAS					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
nonlinear	800	2.011 (0.104)	0.996 (0.035)	0.997 (0.025)	0.919 (0.008)
	1000	2.002 (0.045)	0.999 (0.015)	1.000 (0.011)	0.919 (0.010)
	1200	2.000 (0.000)	1.000 (0.000)	1.000 (0.000)	0.919 (0.009)
riskRFE					
scenario	n	size (sd)	TDR (sd)	MCC (sd)	CCR (sd)
nonlinear	800	3.062 (0.949)	0.713 (0.212)	0.782 (0.171)	0.878 (0.027)
	1000	2.802 (0.848)	0.772 (0.208)	0.829 (0.163)	0.883 (0.028)
	1200	2.561 (0.709)	0.833 (0.192)	0.878 (0.144)	0.894 (0.021)

APPENDIX E

Proof of Algorithm III.1 for Chapter III

According to Algorithm III.1, at iteration k ($k \geq 2$), $(\boldsymbol{\omega}^{(k)}, b^{(k)}, \boldsymbol{\xi}^{(k)})$ is the optimal solution of the optimization problem defined in (3.2). First we prove that $(\boldsymbol{\omega}^{(k-1)}, b^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$ is a feasible solution of (3.2). It immediately follows that $h(\boldsymbol{\omega}^{(k)}, \boldsymbol{\xi}^{(k)}) \leq h(\boldsymbol{\omega}^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$ due to the fact that $(\boldsymbol{\omega}^{(k)}, b^{(k)}, \boldsymbol{\xi}^{(k)})$ is the optimal solution for (3.2). The equality holds when $h(\boldsymbol{\omega}^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$ reaches the minimum. To proceed, it is sufficient to show that the following conditions in (3.2) hold,

$$\begin{aligned}
&A_i(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1 - \xi_i^{(k-1)}, i \in \mathcal{S}_1, \\
&A_i(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1 - \widetilde{\xi}_{li}^{(k-1)}, i \in \tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}, \\
&y_{li}^{(k-1)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1 - \check{\xi}_{li}^{(k-1)}, i \in \tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}.
\end{aligned} \tag{E.1}$$

Note that at iteration $k-1$, $(\boldsymbol{\omega}^{(k-1)}, b^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$ is the solution of problem (3.2)

with $k-1$ replaced by $k-2$,

$$\begin{aligned}
\min_{\boldsymbol{\omega},b,\boldsymbol{\xi}} \quad & \frac{1}{2}||\boldsymbol{\omega}||^2 + C\left[\sum_{i\in\mathcal{S}_1}\frac{B_{1i}}{p_i}\xi_i + \sum_{l=2}^{m}\left\{\lambda_l \sum_{i\in\tilde{\mathcal{S}}_l^{(k)}}\frac{B_{li}}{p_i}\widetilde{\xi}_{li} + (1-\lambda_l)\sum_{i\in\tilde{\mathcal{S}}_l^{(k)}}\frac{\overline{B_l}}{p}\check{\xi}_{li}\right\}\right] \\
\text{s.t.} \quad & A_i(\boldsymbol{\omega}^T\boldsymbol{X}_i + b) \geq 1-\xi_i \text{ and } \xi_i \geq 0,\ i\in\mathcal{S}_1, \\
& A_i(\boldsymbol{\omega}^T\boldsymbol{X}_i + b) \geq 1-\widetilde{\xi}_{li} \text{ and } \widetilde{\xi}_{li} \geq 0,\ i\in\tilde{\mathcal{S}}_l^{(k-1)},\ l\in\mathcal{L}, \\
& y_{li}^{(k-2)}(\boldsymbol{\omega}^T\boldsymbol{X}_i + b) \geq 1-\check{\xi}_{li} \text{ and } \check{\xi}_{li} \geq 0,\ i\in\tilde{\mathcal{S}}_l^{(k-1)},\ l\in\mathcal{L}.
\end{aligned} \tag{E.2}$$

Therefore, we have

$$\begin{aligned}
& A_i(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1-\xi_i^{(k-1)},\ i\in\mathcal{S}_1 \\
& A_i(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1-\widetilde{\xi}_{li}^{(k-1)},\ i\in\tilde{\mathcal{S}}_l^{(k-1)},\ l\in\mathcal{L}.
\end{aligned} \tag{E.3}$$

The constraints involving treatment A in (E.1) are satisfied by $(\boldsymbol{\omega}^{(k-1)}, b^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$. In addition, we have

$$y_{li}^{(k-2)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1-\check{\xi}_{li}^{(k-1)},\ i\in\tilde{\mathcal{S}}_l^{(k-1)},\ l\in\mathcal{L}. \tag{E.4}$$

To prove the constraints involving $y_{li}^{(k-1)}$ in (E.1), we consider two scenarios between predicted labels $y_i^{(k-1)}$ and $y_i^{(k-2)}$ for subject $i\in\bigcup_{l=2}^{m}\tilde{\mathcal{S}}_l$. First, if $y_{li}^{(k-1)} = y_{li}^{(k-2)}$ is true, then we have

$$y_{li}^{(k-1)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) = y_{li}^{(k-2)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1-\check{\xi}_{li}^{(k-1)}. \tag{E.5}$$

Second, if $y_{li}^{(k-1)} \neq y_{li}^{(k-2)}$ is true, then

$$y_{li}^{(k-1)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) > y_{li}^{(k-2)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1-\check{\xi}_{li}^{(k-1)}. \tag{E.6}$$

Since labels are binary and can only take values from $\{-1,1\}$, we must have

$y_{li}^{(k-1)} = -y_{li}^{(k-2)}$ if $y_{li}^{(k-1)} \neq y_{li}^{(k-2)}$. Also, from the definition of $y_i^{(k-1)}$, we have $y_i^{(k-1)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) > 0$. Therefore, $y_{li}^{(k-2)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) < 0$.

Combines the above two cases, we obtain

$$y_{li}^{(k-1)}(\boldsymbol{\omega}^{(k-1)T}\boldsymbol{X}_i + b^{(k-1)}) \geq 1 - \check{\xi}_{li}^{(k-1)}, i \in \tilde{\mathcal{S}}_l^{(k)}, l \in \mathcal{L}, \tag{E.7}$$

where the equality only occurs when the algorithm converges with the predicted labels $y_i, i \in \bigcup_{l=2}^{m} \tilde{\mathcal{S}}_l$ stopping changing between two adjacent iterations $k-1$ and $k-2$.

Thus, all constraints in (E.3) have been proved to hold. This implies that all constraints in (3.2) hold. This completes the proof of the descending property $h(\boldsymbol{\omega}^{(k)}, \boldsymbol{\xi}^{(k)}) < h(\boldsymbol{\omega}^{(k-1)}, \boldsymbol{\xi}^{(k-1)})$ when $k \geq 2$. □

APPENDIX F

Additional Simulation Results for Chapter III

If the proposed parameter tuning method based on the minimization of SSE works well, we should see that the selected $(C, \boldsymbol{\lambda})$ gives an increasing prediction accuracy for the training dataset, a decreasing objective function, and a decreasing SSE. Our experience from extensive simulation studies suggests that Algorithm III.1 converges fast within typically 20 iterations. Thus, we run the algorithm over 20 iterations to see the iteration traces. Figure F.1 illustrates the trajectories of the evaluation criteria over the 20 iterations for the different scenarios. The values shown in each figure are the differences of the evaluation criteria at iteration k versus $k-1$, where $k = 2, \ldots, 20$. In Figure F.1, the increasing of the prediction accuracy, as well as the decreasing of the objective function and SSE are evident as fully expected. Table F.1 summarizes the resulting SSE values corresponding to methods M1-M4 discussed in Section 3.4. As is shown, M1 always gives the smallest SSE. The computation time (in second) of methods M1-M4 is listed in Table F.2, showing that each iteration of our method M1 takes similar running time compared with the existing OWL M4.

Table F.1: SSE (mean (sd)) based on SS-learning with proposed tuning, SS-learning with oracle tuning, OWL on $\mathcal{S}_1$, and OWL on $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$.

	SS-learning (tune)	SS-learning (oracle)	OWL in $\mathcal{S}_1$	OWL on $\mathcal{S}$
$\beta_3 = 0.1$	414.5 (32.3)	424.4 (35.5)	446.78 (42.1)	446.4 (39.9)
$\beta_3 = 0.5$	428.8 (36.1)	435.1 (38.2)	508.2 (51.0)	459.7 (47.0)
$\beta_3 = 1.0$	459.8 (48.2)	467.3 (50.5)	704.2 (86.4)	495.2 (65.3)
$\beta_3 = 3.0$	712.2 (175.8)	743.0 (173.7)	2801.0 (571.6)	863.1 (211.4)
$\sigma_3 = 0.1$	49.5 (12.2)	52.3 (12.7)	144.8 (39.0)	79.5 (37.4)
$\sigma_3 = 0.5$	144.5 (17.5)	148.5 (18.7)	233.0 (40.8)	173.9 (37.1)
$\sigma_3 = 1.0$	430.8 (36.4)	437.0 (37.4)	508.7 (46.5)	460.5 (44.4)
$\sigma_3 = 2.0$	1529.8 (149.6)	1550.4 (105.7)	1611.3 (112.1)	1578.7 (111.0)
$\frac{n_1}{n} = 0.04$	428.8 (48.4)	436.7 (38.2)	509.8 (51.0)	461.1 (45.4)
$\frac{n_1}{n} = 0.20$	536.9 (90.8)	549.6 (93.8)	696.4 (153.4)	613.4 (115.3)
$\frac{n_1}{n} = 0.36$	623.6 (119.5)	638.0 (121.4)	771.3 (168.1)	713.9 (145.3)
$\frac{n_1}{n} = 0.50$	674.7 (149.6)	696.0 (155.9)	807.7 (180.2)	775.2 (181.2)

Table F.2: Computation time in second (mean (sd)) based on SS-learning with proposed tuning, OWL on $\mathcal{S}_1$, and OWL on $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$.

	SS-learning (tune)	OWL on $\mathcal{S}_1$	OWL on $\mathcal{S}$
$\beta_3 = 0.1$	0.34 (0.06)	0.08 (0.01)	0.36 (0.04)
$\beta_3 = 0.5$	0.34 (0.05)	0.07 (0.01)	0.36 (0.04)
$\beta_3 = 1.0$	0.34 (0.05)	0.07 (0.01)	0.36 (0.04)
$\beta_3 = 3.0$	0.33 (0.05)	0.07 (0.01)	0.37 (0.04)
$\sigma_3 = 0.1$	0.33 (0.05)	0.08 (0.01)	0.35 (0.04)
$\sigma_3 = 0.5$	0.33 (0.05)	0.07 (0.01)	0.35 (0.03)
$\sigma_3 = 1.0$	0.34 (0.05)	0.07 (0.01)	0.36 (0.04)
$\sigma_3 = 2.0$	0.35 (0.06)	0.07 (0.01)	0.37 (0.04)
$n_1/n = 0.04$	0.34 (0.06)	0.08 (0.01)	0.36 (0.04)
$n_1/n = 0.20$	0.27 (0.05)	0.08 (0.01)	0.36 (0.04)
$n_1/n = 0.36$	0.19 (0.03)	0.11 (0.01)	0.36 (0.04)
$n_1/n = 0.50$	0.13 (0.02)	0.15 (0.01)	0.36 (0.04)

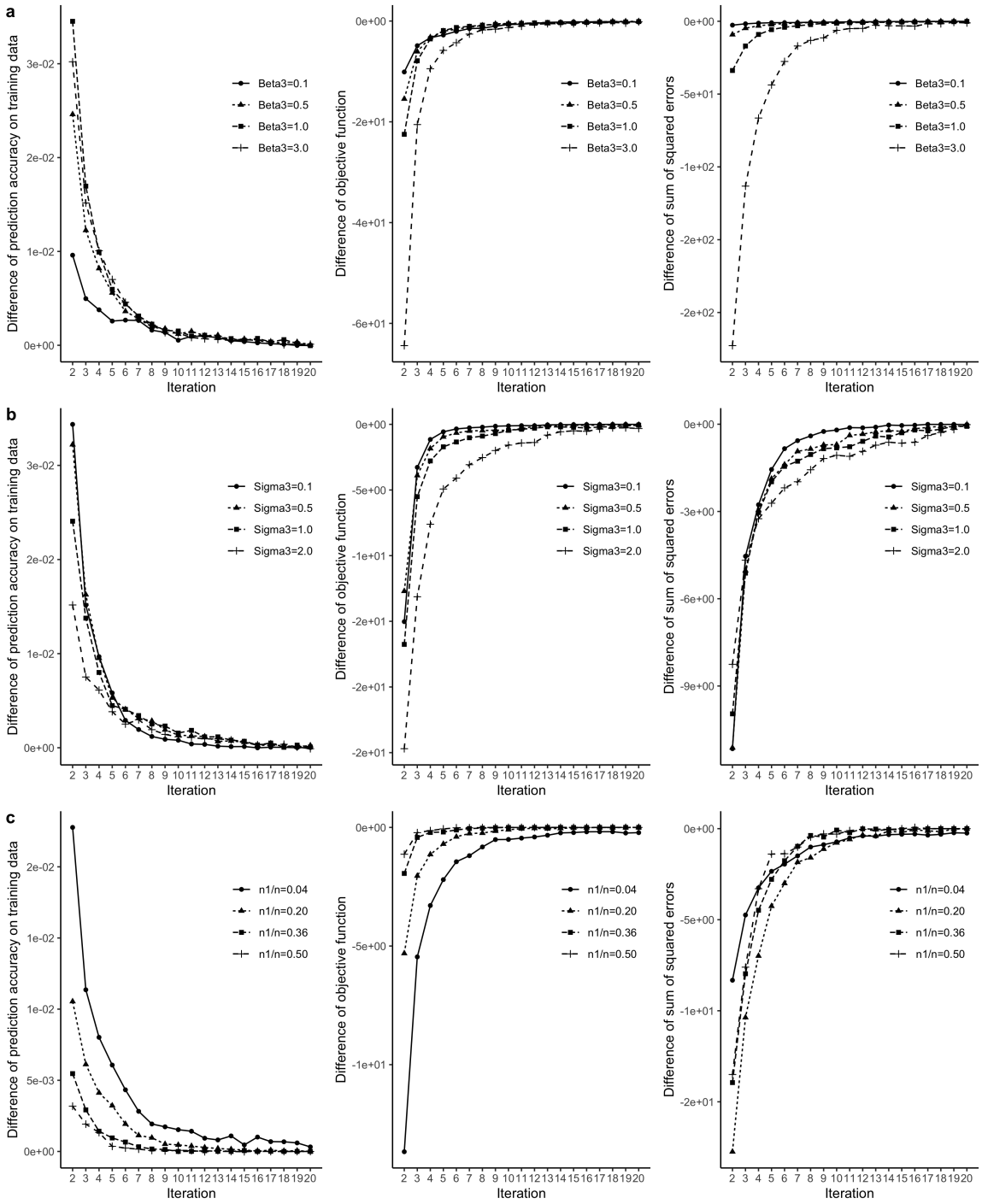

Figure F.1: Convergence analysis under (a) varying β_3, (b) varying σ_3, and (c) varying n_1/n.

APPENDIX G

Nonlinear Simulation Experiment for Chapter III

We here report another simulation experiment, which is designed with two data sources: $\mathcal{S}_1$ contains high quality benefit B_1 and $\mathcal{S}_2$ contains low quality benefit B_2. The feature vector $\boldsymbol{X}$ is generated the same way as specified in Section 3.4. Now, we assume an observational study with treatment $A = 1$ allocated to subjects with probability $p_A = \exp(\eta_A)/(1+\exp(\eta_A))$, where $\eta_A = 0.4+0.5X_1+0.5X_3-X_5-X_7+0.2X_9$. Logistic regression with all the main effects of $\boldsymbol{X}$ is used to estimate the propensity score $P(A \mid \boldsymbol{X})$. The underlying true decision function f is specified as $f(\boldsymbol{X}) = \exp(X_1^2) - \exp(X_2^2)$. The observed benefit B_1 is generated from a normal distribution with mean $\mu_1 = 0.01 + 0.02\sin(3X_4) + 3.0Af(\boldsymbol{X})$ and standard deviation $\sigma_1 = 0.1$, the observed benefit B_2 is generated from a normal distribution with mean $\mu_2 = 0.1 + 0.2\sin(3X_4) + \beta_2 Af(\boldsymbol{X})$ and standard deviation σ_2. Let the total sample size $n = n_1 + n_2$, where $n_1 = |\mathcal{S}_1|$ and $n_2 = |\mathcal{S}_2|$. In order to evaluate the performance of SS-learning under varying relevance of B_1 and B_2, we consider the following scenarios: (d) $\beta_2 \in \{0.1, 0.5, 1.0, 3.0\}, \sigma_2 = 1.0, n_1 = 100, n_2 = 400$; (e) $\beta_2 = 0.5, \sigma_2 \in \{0.1, 0.5, 1.0, 2.0\}, n_1 = 100, n_2 = 400$; (f) $\beta_2 = 0.5, \sigma_2 = 1.0, n_1 \in \{20, 100, 180, 250\}, n_2 \in \{480, 400, 320, 250\}$. The optimization function is solved using the RBF kernel. We set tuning parameters $C \in \{1/128, 1/64, \ldots, 1/2, 1\}, \gamma \in$

$\{1/128, 1/64, \ldots, 1/2, 1\}$ and weighting parameter $\lambda \in \{0, 0.1, \ldots, 0.9, 1.0\}$ for parameter selection. Two different regression models are fitted for $g_l(\boldsymbol{X}), l = 1, 2$ in parameter tuning, i) a simple linear regression with all the main effects of $\boldsymbol{X}$, and ii) a GAM with B-spline basis functions for $\boldsymbol{X}$ (20 basis functions for each variable).

We compare the prediction accuracy on the training and validation datasets given by the four different methods: (M1) SS-learning with tuning according to the minimization of SSE; (M2) oracle SS-learning with optimal tuning according to the highest prediction accuracy on $\mathcal{S}_1 \cup \mathcal{S}_2$; (M3) OWL with 5-fold cross-validation based tuning of C in SVM on $\mathcal{S}_1$ alone; (M4) OWL with 5-fold cross-validation based tuning of C in SVM on $\mathcal{S}_1 \cup \mathcal{S}_2$. Similar to the results illustrated in Section 3.4, Figure G.1 and G.2 show that the oracle SS-learning M2 always gives the highest prediction accuracy. The M1 method gives a slightly lower prediction accuracy compared to M2. The prediction accuracy given by M4 is constantly lower than M1 and M2. M3 usually obtains the smallest prediction accuracy, except for the cases when the clinical relevance of B_2 is too small (i.e. $\beta_2 = 0.1$) or when B_2 has too much noise (i.e. $\sigma_2 = 2.0$), in which M4 gives the smallest prediction accuracy. It is notable that the proposed M2 gives slightly lower prediction accuracy than M4 in the scenario of $n_1/n = 0.04$ when parameter tuning is performed using linear regression, which is not observed when parameter tuning is carried out by GAM. This stresses the importance of parameter tuning using the correct model that better characterizes the relationship of the benefit value and the covariates. Table G.1 lists the selected λ_2 values. Generally speaking, λ_2 increases with increasing β_2, decreases with increasing σ_2 and n_1/n, with exceptions when $\sigma_2 = 2.0$ and $n_1/n = 0.50$, indicating the difficulty of tuning λ_2 when B_2 has too much noise and $\mathcal{S}_1$ has large sharing in the whole dataset.

Table G.1: Selected weighting values λ_2 (mean (sd)) based on SS-learning using the proposed tuning method by linear regression and GAM with single low-quality dataset.

scenario	λ_2 (linear)	λ_2 (GAM)
$\beta_2 = 0.1$	0.59 (0.38)	0.55 (0.39)
$\beta_2 = 0.5$	0.86 (0.27)	0.86 (0.27)
$\beta_2 = 1.0$	0.94 (0.16)	0.93 (0.18)
$\beta_2 = 3.0$	0.98 (0.07)	0.97 (0.08)
$\sigma_2 = 0.1$	0.97 (0.09)	0.96 (0.11)
$\sigma_2 = 0.5$	0.92 (0.18)	0.92 (0.19)
$\sigma_2 = 1.0$	0.86 (0.25)	0.84 (0.27)
$\sigma_2 = 2.0$	0.72 (0.37)	0.74 (0.35)
$n_1/n = 0.04$	0.98 (0.07)	0.92 (0.16)
$n_1/n = 0.20$	0.88 (0.24)	0.85 (0.27)
$n_1/n = 0.36$	0.84 (0.25)	0.83 (0.26)
$n_1/n = 0.50$	0.81 (0.26)	0.80 (0.25)

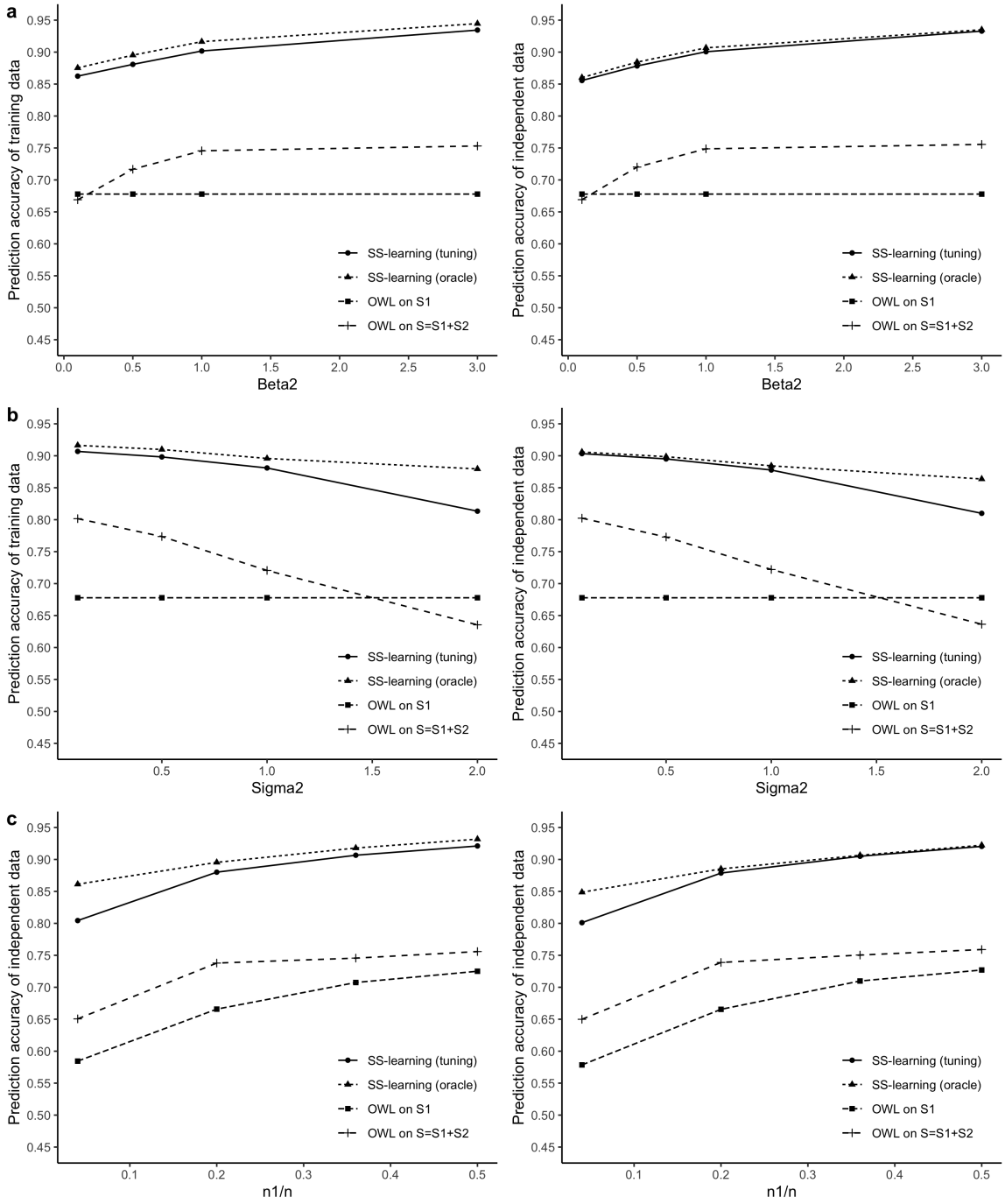

Figure G.1: Comparison of prediction accuracy on the training and validation datasets by different methods with single low-quality dataset using parameter tuning based on GAM under (a) varying β_2, (b) varying σ_2, (c) varying n_1/n.

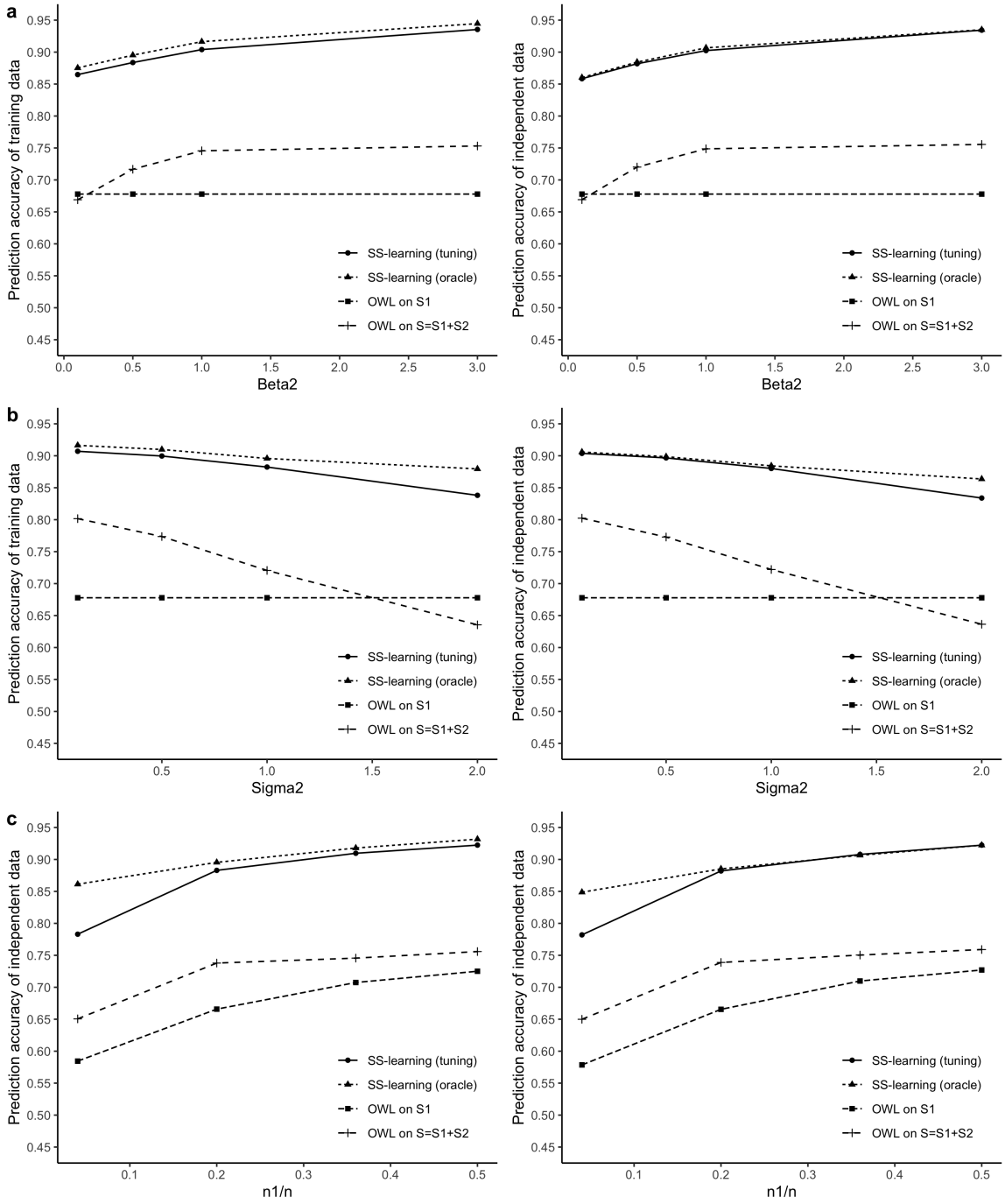

Figure G.2: Comparison of prediction accuracy on the training and validation datasets by different methods with single low-quality dataset using parameter tuning based on linear regression under (a) varying β_2, (b) varying σ_2, (c) varying n_1/n.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abdulsalam, H., D. B. Skillicorn, and P. Martin (2010), Classification using streaming random forests, *IEEE Transactions on knowledge and Data Engineering*, *23*(1), 22–36.

Austin, P. C. (2009), Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples, *Statistics in medicine*, *28*(25), 3083–3107.

Benjamini, Y., and Y. Hochberg (1995), Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Berkane, M. (2012), *Latent variable modeling and applications to causality*, vol. 120, Springer Science & Business Media.

Braun, J. M., et al. (2012), Assessing windows of susceptibility to lead-induced cognitive deficits in mexican children, *Neurotoxicology*, *33*(5), 1040–1047.

Charbonnel, B., D. Matthews, G. Schernthaner, M. Hanefeld, P. Brunetti, and Q. S. Group (2005), A long-term comparison of pioglitazone and gliclazide in patients with type 2 diabetes mellitus: a randomized, double-blind, parallel-group comparison trial, *Diabetic Medicine*, *22*(4), 399–405.

Chen, A., K. N. Dietrich, J. H. Ware, J. Radcliffe, and W. J. Rogan (2005), Iq and blood lead from 2 to 7 years of age: are the effects in older children the residual of high blood lead concentrations in 2-year-olds?, *Environmental Health Perspectives*, *113*(5), 597–601.

Chen, G., D. Zeng, and M. R. Kosorok (2016), Personalized dose finding using outcome weighted learning, *Journal of the American Statistical Association*, *111*(516), 1509–1521.

Collins, F. S., and H. Varmus (2015), A new initiative on precision medicine, *New England journal of medicine*, *372*(9), 793–795.

Cortes, C., and V. Vapnik (1995), Support-vector networks, *Machine learning*, *20*(3), 273–297.

Council, N. R., et al. (2011), *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*, National Academies Press.

Dasgupta, S., Y. Goldberg, M. R. Kosorok, et al. (2019), Feature elimination in kernel machines in moderately high dimensions, *The Annals of Statistics*, *47*(1), 497–526.

Ettinger, A. S., M. M. Téllez-Rojo, C. Amarasiriwardena, K. E. Peterson, J. Schwartz, A. Aro, H. Hu, and M. Hernández-Avila (2005), Influence of maternal bone lead burden and calcium intake on levels of lead in breast milk over the course of lactation, *American Journal of Epidemiology*, *163*(1), 48–56.

Ettinger, A. S., H. Hu, and M. Hernandez-Avila (2007), Dietary calcium supplementation to lower blood lead levels in pregnancy and lactation, *The Journal of nutritional biochemistry*, *18*(3), 172–178.

Fan, A., W. Lu, and R. Song (2016), Sequential advantage selection for optimal treatment regime, *The annals of applied statistics*, *10*(1), 32.

Glynn, A. N., and K. M. Quinn (2010), An introduction to the augmented inverse propensity weighted estimator, *Political analysis*, pp. 36–56.

Gulson, B., K. Mahaffey, K. Mizon, M. Korsch, M. Cameron, and G. Vimpani (1995), Contribution of tissue lead to blood lead in adult female subjects based on stable lead isotope methods., *The Journal of laboratory and clinical medicine*, *125*(6), 703–712.

Gulson, B. L., K. J. Mizon, M. J. Korsch, J. M. Palmer, and J. B. Donnelly (2003), Mobilization of lead from human bone tissue during pregnancy and lactation—a summary of long-term research, *Science of the Total Environment*, *303*(1-2), 79–104.

Gulson, B. L., K. J. Mizon, J. M. Palmer, M. J. Korsch, A. J. Taylor, and K. R. Mahaffey (2004), Blood lead changes during pregnancy and postpartum with calcium supplementation, *Environmental health perspectives*, *112*(15), 1499–1507.

Gunter, L., J. Zhu, and S. Murphy (2011), Variable selection for qualitative interactions, *Statistical methodology*, *8*(1), 42–55.

Hornung, R. W., B. P. Lanphear, and K. N. Dietrich (2009), Age of greatest susceptibility to childhood lead exposure: a new statistical approach, *Environmental Health Perspectives*, *117*(8), 1309–1312.

Hu, H., and M. Hernandez-Avila (2002), Invited commentary: lead, bones, women, and pregnancy—the poison within?, *American journal of epidemiology*, *156*(12), 1088–1091.

Hu, H., et al. (2006), Fetal lead exposure at each stage of pregnancy as a predictor of infant mental development, *Environmental health perspectives*, *114*(11), 1730–1735.

Janakiraman, V., A. Ettinger, A. Mercado-Garcia, H. Hu, and M. Hernandez-Avila (2003), Calcium supplements and bone resorption in pregnancy: a randomized crossover trial, *American journal of preventive medicine*, *24*(3), 260–264.

Jones, D. T., et al. (2019), Molecular characteristics and therapeutic vulnerabilities across paediatric solid tumours, *Nature Reviews Cancer*, *19*(8), 420–438.

Little, R. J. (2008), Selection and pattern-mixture models, *Longitudinal data analysis*, pp. 409–431.

Liu, S., and V. E. Johnson (2016), A robust bayesian dose-finding design for phase i/ii clinical trials, *Biostatistics*, *17*(2), 249–263.

Lu, W., H. H. Zhang, and D. Zeng (2013), Variable selection for optimal treatment decision, *Statistical methods in medical research*, *22*(5), 493–504.

Lu, Y.-F., D. B. Goldstein, M. Angrist, and G. Cavalleri (2014), Personalized medicine and human genetic diversity, *Cold Spring Harbor perspectives in medicine*, *4*(9), a008,581.

Mendelsohn, A. L., B. P. Dreyer, A. H. Fierman, C. M. Rosen, L. A. Legano, H. A. Kruger, S. W. Lim, and C. D. Courtlandt (1998), Low-level lead exposure and behavior in early childhood, *Pediatrics*, *101*(3), e10–e10.

Molnar, F. J., B. Hutton, and D. Fergusson (2008), Does analysis using "last observation carried forward" introduce bias in dementia research?, *Cmaj*, *179*(8), 751–753.

Mu, X., F. Zhu, J. Du, E.-P. Lim, and Z.-H. Zhou (2017), Streaming classification with emerging new class by class matrix sketching, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31.

Murphy, S. A. (2003), Optimal dynamic treatment regimes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(2), 331–355.

Murphy, S. A. (2005), An experimental design for the development of adaptive treatment strategies, *Statistics in medicine*, *24*(10), 1455–1481.

Murphy, S. A., M. J. van der Laan, J. M. Robins, and C. P. P. R. Group (2001), Marginal mean models for dynamic regimes, *Journal of the American Statistical Association*, *96*(456), 1410–1423.

Nair, L. R., S. D. Shetty, and S. D. Shetty (2018), Applying spark based machine learning model on streaming big data for health status prediction, *Computers & Electrical Engineering*, *65*, 393–399.

Niculescu, A., et al. (2019), Towards precision medicine for pain: diagnostic biomarkers and repurposed drugs, *Molecular psychiatry*, *24*(4), 501–522.

Norman, J., D. Politz, and L. Politz (2009), Hyperparathyroidism during pregnancy and the effect of rising calcium on pregnancy loss: a call for earlier intervention, *Clinical endocrinology*, *71*(1), 104–109.

Pencina, M. J., R. B. D'Agostino, and R. S. Vasan (2010), Statistical methods for assessment of added usefulness of new biomarkers, *Clinical Chemistry and Laboratory Medicine (CCLM)*, *48*(12), 1703–1711.

Pepe, M. S., H. Janes, and C. I. Li (2014), Net risk reclassification p values: valid or misleading?, *Journal of the National Cancer Institute*, *106*(4), dju041.

Perng, W., et al. (2019), Early life exposure in mexico to environmental toxicants (element) project, *BMJ open*, *9*(8), e030,427.

Qian, M., and S. A. Murphy (2011), Performance guarantees for individualized treatment rules, *Annals of statistics*, *39*(2), 1180.

Street, W. N., and Y. Kim (2001), A streaming ensemble algorithm (sea) for large-scale classification, in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 377–382.

Téllez-Rojo, M., H. Lamadrid-Figueroa, A. Mercado-García, K. Peterson, D. Bellinger, A. Ettinger, M. Solano-González, M. Heránndez-Ávila, and H. Hu (2006a), A randomized controlled trial of calcium supplementation to reduce blood lead levels (and fetal lead exposure) in pregnant women, *Epidemiology*, *17*(6), S123.

Téllez-Rojo, M. M., M. Hernández-Avila, H. Lamadrid-Figueroa, D. Smith, L. Hernández-Cadena, A. Mercado, A. Aro, J. Schwartz, and H. Hu (2004), Impact of bone lead and bone resorption on plasma and whole blood lead levels during pregnancy, *American journal of epidemiology*, *160*(7), 668–678.

Téllez-Rojo, M. M., D. C. Bellinger, C. Arroyo-Quiroz, H. Lamadrid-Figueroa, A. Mercado-Garcia, L. Schnaas-Arrieta, R. O. Wright, M. Hérnandez-Avila, and H. Hu (2006b), Longitudinal associations between blood lead concentrations lower than 10 μg/dl and neurobehavioral development in environmentally exposed children in mexico city, *Pediatrics*, *118*(2), e323–e330.

Thomas, L. E., F. Li, and M. J. Pencina (2020), Overlap weighting: a propensity score method that mimics attributes of a randomized clinical trial, *Jama*, *323*(23), 2417–2418.

van Buuren, S., and K. Groothuis-Oudshoorn (2011), mice: Multivariate imputation by chained equations in r, *Journal of Statistical Software*, *45*(3), 1–67.

Vickers, A. J., and E. B. Elkin (2006), Decision curve analysis: a novel method for evaluating prediction models, *Medical Decision Making*, *26*(6), 565–574.

Vickers, A. J., and M. Pepe (2014), Does the net reclassification improvement help us evaluate models and markers?, *Annals of internal medicine*, *160*(2), 136–137.

Wadhwa, P. D., C. Buss, S. Entringer, and J. M. Swanson (2009), Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms, in *Seminars in reproductive medicine*, vol. 27, pp. 358–368, © Thieme Medical Publishers.

Wasserman, G. A., B. Staghezza-Jaramillo, P. Shrout, D. Popovac, and J. Graziano (1998), The effect of lead exposure on behavior problems in preschool children., *American Journal of Public Health*, *88*(3), 481–486.

Wasserman, G. A., P. Factor-Litvak, X. Liu, A. C. Todd, J. K. Kline, V. Slavkovich, D. Popovac, and J. H. Graziano (2003), The relationship between blood lead, bone lead and child intelligence, *Child Neuropsychology*, *9*(1), 22–34.

Watkins, C. J. C. H. (1989), Learning from delayed rewards, Ph.D. thesis, King's College, Cambridge.

Wu, P., D. Zeng, and Y. Wang (2020), Matched learning for optimizing individualized treatment strategies using electronic health records, *Journal of the American Statistical Association*, *115*(529), 380–392.

Yau, T. O. (2019), Precision treatment in colorectal cancer: Now and the future, *JGH Open*, *3*(5), 361–369.

Zhang, A., et al. (2012), Association between prenatal lead exposure and blood pressure in children, *Environmental Health Perspectives*, *120*(3), 445–450.

Zhao, Y., D. Zeng, A. J. Rush, and M. R. Kosorok (2012), Estimating individualized treatment rules using outcome weighted learning, *Journal of the American Statistical Association*, *107*(499), 1106–1118.

Zhou, X., N. Mayer-Hamblett, U. Khan, and M. R. Kosorok (2017), Residual weighted learning for estimating individualized treatment rules, *Journal of the American Statistical Association*, *112*(517), 169–187.

Zhu, L., W. Lu, M. R. Kosorok, and R. Song (2020), Kernel assisted learning for personalized dose finding, in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 56–65.